

**Cristhian Alberto Celestino Cortez**

**AVALIAÇÃO DE UM NOVO MÉTODO DE GERAÇÃO AUTOMÁTICA  
DE IMAGENS PANÔRAMICAS DE FUNDO DE UMA SEQUÊNCIA  
DE VÍDEO**

Dissertação apresentada como requisito parcial à obtenção do grau de Mestre, no Programa de Pós-Graduação em Engenharia Elétrica, Setor de Tecnologia, Universidade Federal do Paraná – UFPR.

Orientador: Eduardo Parente Ribeiro, Dr.  
Co-orientador: Marcus Vinicius Lamar, Dr.

Curitiba  
2005

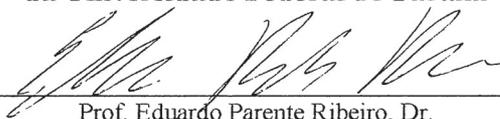
# UM NOVO MÉTODO DE GERAÇÃO AUTOMÁTICA DE IMAGENS PANORÂMICAS DE FUNDO DE UMA SEQÜÊNCIA DE VÍDEO

CRISTHIAN ALBERTO CELESTINO CORTEZ

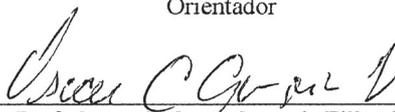
Dissertação aprovada como requisito parcial para obtenção do grau de Mestre no

**Programa de Pós-Graduação em Engenharia Elétrica**

**da Universidade Federal do Paraná**



Prof. Eduardo Parente Ribeiro, Dr.  
Orientador

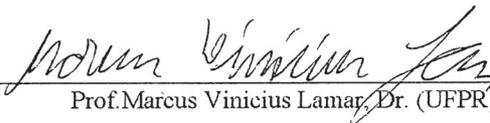


Prof. Oscar da Costa Gouveia Filho, Dr.  
Coordenador do Programa de Pós-Graduação em Engenharia Elétrica

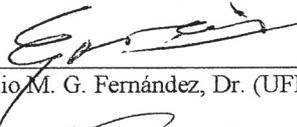
## **Banca Examinadora**



Prof. Eduardo Parente Ribeiro, Dr. (UFPR)  
Presidente



Prof. Marcus Vinicius Lamar, Dr. (UFPR)



Prof. Evélio M. G. Fernández, Dr. (UFPR)



Prof. Marcelo Eduardo Pellenz, Dr. (PUC-PR)

Curitiba, 10 de março de 2005

## **Agradecimentos**

Ao Prof. Dr. Eduardo Parente Ribeiro por seus ensinamentos e conselhos.

Ao Prof. Dr. Marcus Vinicius Lamar que contribuiu para a elaboração desse trabalho através de seu conhecimento e de suas sugestões.

À minha família, meu pai Aldemir Del Cortez, minha mãe Teresa Celestino Cortez, minha irmã Teresa Cristina Celestino Cortez, meu sobrinho Gabriel, meu cunhado Miguel e meus tios e tias, pelo apoio e pelas palavras de incentivo.

A todos aqueles que me ajudaram e forneceram discussões e dicas valiosas. Em especial, aos amigos Fernando Nadal, Marcos Vinicius de Paula Pinto, André Augusto Mariano, Cícero Hildenberg, Marcelo Nascimento, Rafael Palma, Michelle Foltran, Alberto Natory, Salmo Pustilnick, Marcelo Naegele, Sinésio Júlio Barberini, Walkelly, Igor Gavriloff e Fernando César de Oliveira.

À Capes pela bolsa .

# Sumário

<b>Agradecimentos</b> .....	<b>ii</b>
<b>Sumário</b> .....	<b>iii</b>
<b>Lista de figuras</b> .....	<b>iv</b>
<b>Lista de tabelas</b> .....	<b>v</b>
<b>Lista de acrônimos</b> .....	<b>vi</b>
<b>Resumo</b> .....	<b>vii</b>
<b>Abstract</b> .....	<b>viii</b>
<b>1. Introdução</b> .....	<b>1</b>
<b>2. Métodos Aplicados à Estimação de Mosaico</b> .....	<b>8</b>
<b>3. Metodologia da estimação de mosaico</b> .....	<b>16</b>
3.1. Método de Estimação de Movimento Global.....	16
3.2. Pontos específicos do método de estimação do movimento global.....	21
3.3. Método de composição do mosaico.....	36
3.4. Padrões e critérios de desempenho.....	39
3.5. Considerações finais do capítulo .....	41
<b>4. Resultados</b> .....	<b>42</b>
4.1. Mosaicos.....	43
4.2. Relação entre a qualidade da compensação, zonas de atividade de movimento e padrões de compensação .....	57
4.3. Comparação entre os vídeos naturais .....	68
4.4. Influência da interpolação na qualidade do mosaico.....	71
4.5. Avaliação do Método de Estimação dos Parâmetros de Movimento Global.....	81
4.6. Comparação com outros métodos .....	95
<b>5. Conclusão</b> .....	<b>98</b>
5.1. Trabalhos futuros .....	99
<b>Referências Bibliográficas</b> .....	<b>100</b>

## Lista de figuras

Figura 1- Objeto e fundo.....	2
Figura 2- Composição da cena. ....	3
Figura 3- Fluxograma do método de estimação de movimento global.....	18
Figura 4- Vetores de movimento de regiões homogêneas.....	26
Figura 5- Classificação dos vetores de movimento. ....	28
Figura 6- Exemplo de histograma, para a obtenção da moda .....	29
Figura 7- Extração dos vetores de movimento dos objetos.....	33
Figura 8- Acúmulo de erros nos parâmetros de rotação.....	35
Figura 9-Alguns quadros da seqüência coast_guard. ....	43
Figura 10- Mosaico de 100 quadros da seqüência coast guard. ....	44
Figura 11-Alguns quadros da seqüência Stefan. ....	44
Figura 12- Mosaico de 100 quadros da seqüência Stefan. ....	45
Figura 13- Alguns quadros da seqüência bus. ....	46
Figura 14- Mosaico de 100 quadros da seqüência Bus. ....	47
Figura 15- Alguns quadros da seqüência Lenna_zoom.....	47
Figura 16- Mosaico da seqüência Lennazoom.....	48
Figura 17- Alguns quadros da seqüência Luana_zoom.....	48
Figura 18-Alguns quadros da seqüência peixe_zoom. ....	48
Figura 19- Mosaico da seqüência Luana_zoom.....	49
Figura 20- Mosaico da seqüência peixe_zoom. ....	49
Figura 21-Alguns quadros da seqüência Lenna_rotação. ....	50
Figura 22- Trajetória do ponto central do primeiro quadro da seqüência Lenna_rotação. ....	50
Figura 23- Mosaico da seqüência Lenna_rotação. ....	51
Figura 24- Trajetória do ponto central do primeiro quadro da seqüência Luana_rotação. ....	51
Figura 25-Alguns quadros da seqüência Luana_rotação. ....	52
Figura 26- Mosaico da seqüência Luana_rotação. ....	52
Figura 27- Mosaico da seqüência peixe_rotação.....	53
Figura 28-Alguns quadros da seqüência peixe_rotação.....	53
Figura 29- Trajetória do ponto central do primeiro quadro da seqüência peixe_rotação. ....	54
Figura 30- Mosaico da seqüência Lenna_completa. ....	54
Figura 31- Alguns quadros da seqüência Lenna_completa. ....	54
Figura 32- Mosaico da seqüência peixe_translação.....	55
Figura 33-Alguns quadros da seqüência peixe_translação.....	55
Figura 34-Trajetoira do ponto central do primeiro quadro da seqüência peixe_translação.....	56
Figura 35- Seqüência bus, comparação entre os métodos.....	58
Figura 36- Seqüência Stefan, comparação entre os métodos.....	59
Figura 37- Seqüência coast_guard, comparação entre os métodos. ....	60
Figura 38- Seqüência Lenna_completa, comparação entre os métodos.....	61

Figura 39- Seqüência Lenna_ rotação, comparação entre os métodos.....	62
Figura 40- Seqüência Luana_ rotação, comparação entre os métodos.....	63
Figura 41- Seqüência peixe_ rotação, comparação entre os métodos. ....	64
Figura 42- Seqüência Lenna_ zoom, comparação entre os métodos. ....	65
Figura 43- Seqüência Luana_ zoom, comparação entre os métodos. ....	66
Figura 44- Seqüência peixe_ zoom, comparação entre os métodos.....	66
Figura 45- Seqüência peixe_ translação, comparação entre os métodos. ....	67
Figura 46- Comparação entre o PSNR dos vídeos naturais, compensação por mosaico . ....	69
Figura 47- Comparação entre o PSNR dos vídeos naturais, compensação quadro a quadro . ....	69
Figura 48- Comparação entre o PSNR dos vídeos naturais, sem compensação. .	71
Figura 49- Relação entre a interpolação e o PSNR, Lenna_ completa.....	73
Figura 50- Relação entre a interpolação e o PSNR, Lenna_ rotação.....	74
Figura 51- Relação entre a interpolação e o PSNR, peixe_ rotação. ....	75
Figura 52- Relação entre a interpolação e o PSNR, Luana_ rotação.....	76
Figura 53- Relação entre a interpolação e o PSNR, Lenna_ zoom. ....	77
Figura 54- Relação entre a interpolação e o PSNR, Luana_ zoom. ....	78
Figura 55- Relação entre a interpolação e o PSNR, peixe_ zoom.....	79
Figura 56- Relação entre a interpolação e o PSNR, peixe_ translação. ....	80
Figura 57- Relação entre os parâmetros de movimento estimados e os reais I, Lenna_ completo. ....	82
Figura 58- Relação entre os parâmetros de movimento estimados e os reais II, Lenna_ completo. ....	83
Figura 59- Relação entre os parâmetros de movimento estimados e os reais III, Lenna_ completo. ....	84
Figura 60- Relação entre os parâmetros de movimento estimados e os reais I, Lenna_ rotação. ....	85
Figura 61- Relação entre os parâmetros de movimento estimados e os reais II, Lenna_ rotação. ....	86
Figura 62- Relação entre os parâmetros de movimento estimados e os reais III, Lenna_ rotação. ....	87
Figura 63- Relação entre os parâmetros de movimento estimados e os reais I, Lenna_ zoom.....	89
Figura 64- Relação entre os parâmetros de movimento estimados e os reais II, Lenna_ zoom.....	90
Figura 65- Relação entre os parâmetros de movimento estimados e os reais I, peixe_ zoom. ....	91
Figura 66- Relação entre os parâmetros de movimento estimados e os reais II, peixe_ zoom. ....	92
Figura 67- Relação entre os parâmetros de movimento estimados e os reais I, peixe_ translação.....	94

## Lista de tabelas

Tabela 1- Comparação entre os métodos de estimação de mosaico .....	95
---	----

## Lista de acrônimos

<b>Sigla</b>	<b>Definição</b>	<b>Significado</b>
<b>ACV</b>	-	campo de vetores artificial
<b>ACV<sub>x</sub></b>	-	coordenadas em x do campo de vetores artificial
<b>ACV<sub>y</sub></b>	-	coordenadas em y do campo de vetores artificial
<b>BMA</b>	<i>block match algorithm</i>	algoritmo por procura de bloco
<b>CC</b>	<i>cross correlation</i>	correlação cruzada
<b>CIF</b>	<i>common intermediate format</i>	formato intermediário comum
<b>CV</b>	-	campo de vetores
<b>CV<sub>x</sub></b>	-	coordenadas em x do campo de vetores
<b>CV<sub>y</sub></b>	-	coordenadas em y do campo de vetores
<b>DCT</b>	<i>discrete cossine transform</i>	transformada discreta de cossenos
<b>EBSLD</b>	-	escolha baseada na soma dos limiares das dispersões
<b>MAE</b>	<i>mean absolut error</i>	erro médio absoluto
<b>MC-DCT</b>	<i>motion compensation with discrete cossine transform</i>	transformada discreta de cossenos com compensação de movimento
<b>MPEG</b>	<i>movie picture expert group</i>	
<b>MSE</b>	<i>mean square error</i>	erro médio quadrático
<b>PMG</b>	-	parâmetros de movimento global
<b>PSNR</b>	<i>peak signal to noise ratio</i>	relação sinal ruído de pico
<b>RGB</b>	<i>red green blue</i>	vermelho verde e azul
<b>RMS</b>	<i>root mean square</i>	raiz da média quadrática
<b>SLD</b>	-	soma dos limiares das dispersões
<b>SMSE</b>	-	superfície de erros médios quadráticos
<b>SVD</b>	<i>singular value decomposition</i>	decomposição em valor singular
<b>VM</b>	-	vetor de movimento
<b>VO</b>	<i>visual objects</i>	objeto visual
<b>VOB</b>	-	vetores de movimento de objetos

## Resumo

Neste trabalho é proposto um esquema de geração da imagem panorâmica do fundo de uma seqüência de vídeo. A imagem panorâmica do fundo, também conhecida como mosaico ou *sprite*, é gerada a partir da estimação dos parâmetros de movimento global entre o quadro atual e o mosaico. Foi desenvolvido e avaliado um método de estimação de movimento global, para ser aplicado na geração do mosaico. O novo método de estimação do movimento global é baseado no fluxo dos *pixels*, vetores de movimento, que são filtrados para posterior estimação dos parâmetros de movimento global.

**Palavras-chaves:** Mosaico, *sprite* e estimação global de movimento.

## **ABSTRACT**

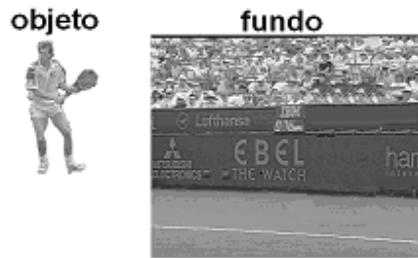
In this work it is proposed an estimation scheme for background panoramic image of video sequence. The background panoramic image also called mosaic or sprite, composed after the global motion parameters estimation between sprite and current frame. A new method of global motion estimation was developed and tested, for sprite generation. The new method of global motion estimation is based on pixels flow and on the filtering of motion vectors for estimation of global motion parameters.

**Keywords** : Mosaic, Sprite and global motion estimation.

# 1.Introdução

As técnicas de compressão de vídeo têm como objetivo reduzir o tamanho dos arquivos através das informações redundantes e explorando as limitações da visão humana. Para diminuir as informações redundantes entre os quadros de vídeo são utilizadas técnicas chamadas de compensações de movimento, que correlacionam regiões entre os quadros de uma seqüência de vídeo. A compensação de movimento é uma ferramenta poderosa, sendo a base de técnicas de compressão como o MPEG-1 [1] e o MPEG-2 [2]. Nos padrões MPEG-1 e MPEG-2 são propostas compensações de movimento local, na qual os blocos de um quadro são correlacionados com os blocos de outro quadro, segundo um critério de similaridade. Dessa forma é necessário transmitir somente a direção da translação do bloco de um quadro em relação ao bloco do outro quadro e a diferença entre os blocos dos dois quadros. Porém, novas técnicas de codificação estão sendo desenvolvidas com o objetivo de reduzir, ainda mais, a quantidade de informação redundante em uma seqüência de vídeo. Algumas dessas novas técnicas de codificação de vídeo fazem parte do padrão MPEG-4 [3].

O padrão MPEG-4, como o MPEG-1 e MPEG-2, apresenta um conjunto de técnicas que aproveitam as redundâncias existentes entre os *pixels* de um mesmo quadro, redundâncias espaciais, e a redundância entre quadros diferentes, redundâncias temporais. No padrão MPEG-4 são encontrados vários tipos de técnicas de predição e compensação de movimento, utilizando a redundância temporal. No padrão MPEG-4, estas técnicas são voltadas para os objetos visuais (VO). Os VO's são as partes que compõem a imagem de um vídeo. Um exemplo de VO é uma pessoa, sem o fundo do quadro, ou o próprio fundo do quadro, como mostra a Figura 1.



**Figura 1-** Objeto e fundo.

A utilização de VO's apresenta como vantagem um melhor aproveitamento da redundância espacial e temporal.

Alguns exemplos de técnicas de predição e compensação de movimento, empregados no padrão MPEG-4, são:

- Estimação e compensação de movimento através de blocos 8x8 e 16x16.
- Compensação de movimento global usando 8 parâmetros de descrição de movimento.
- Compensação de movimento global baseada na transmissão de *sprite* estático.

A imagem panorâmica do fundo, *sprite*, tema deste trabalho, é gerada através de uma seqüência de quadros, de onde são extraídos os *pixels*, que fazem parte do fundo. Os *pixels* do fundo são agregados na imagem panorâmica após uma compensação de movimento. Para quantificar a compensação de movimento, que será empregada no *pixel*, é necessário um estimador de movimento global. O movimento global pode ser descrito por somente 8 parâmetros, os parâmetros de movimento global [3]. Esses parâmetros descrevem o movimento de câmera e são codificados para reconstruir o fundo do quadro original. Com o quadro reconstruído, a partir do mosaico, é fundido o objeto, que é codificado separadamente, para reconstruir o quadro original.

Na nomenclatura adotada pelo padrão MPEG-4, os mosaicos são divididos em duas categorias:

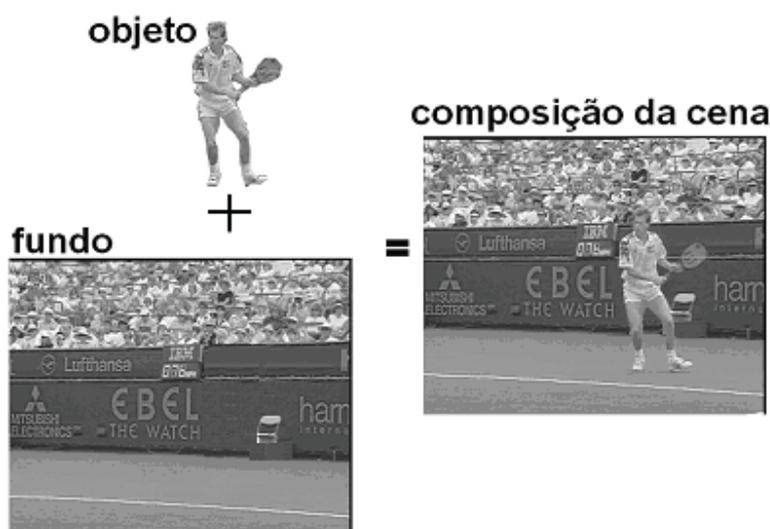
- Dinâmicos.
- Estáticos.

Os estáticos, que serão objeto de estudo deste trabalho, são gerados antes do começo da transmissão do vídeo, não sendo utilizados em aplicações em

tempo real. Já os *sprites* dinâmicos, são gerados durante a execução do vídeo, como referência para as codificações preditivas dos quadros.

Acredita-se que a codificação através do mosaico deve aumentar a taxa de compressão devido à alta redundância espacial presente no mosaico. Porém é necessário que a qualidade do mosaico seja muito boa, para que a reconstrução dos quadros a partir do mosaico seja eficaz.

No padrão MPEG-4 também é prevista a fusão de objetos sintéticos com vídeo natural. Para ocorrer este tipo de fusão, é necessário que o vídeo natural seja codificado em objetos e fundo, VO's, para facilitar a manipulação do vídeo natural, como mostra a Figura 2.



**Figura 2-** Composição da cena.

O processo de estimação do mosaico possui dois pontos críticos. O primeiro é a estimação do movimento global, movimento do fundo da seqüência, com precisão suficiente para não diminuir a qualidade do mosaico. O segundo problema é a segmentação dos objetos que estão na frente do fundo da seqüência de vídeo. Estes dois problemas não têm soluções triviais e têm um desempenho diferente para diferentes vídeos.

Os mosaicos são úteis, em termos de predição de quadros, em situações em que aparecem no quadro novas regiões, impossíveis de serem preditas pelo

quadro anterior ou poucos quadros anteriores [4]. Essas situações são típicas de movimentos de translação, onde por muitas vezes novas regiões aparecem entre quadros subseqüentes. Outras situações em que os mosaicos são úteis são no zoom e em grandes rotações, somente nos casos em que aparecem regiões novas entre os quadros.

Outra utilidade do mosaico provém da separação do quadro de vídeo natural em objetos visuais, para possíveis manipulações [3]. Com esta separação é possível, através de linguagens de descrição de cenas, manipular o vídeo natural da mesma forma que são manipulados os vídeos artificiais. Esta possibilidade transforma o usuário, de um telespectador, que possui poucas opções, em um ente ativo na composição das cenas.

Caso o mosaico possua uma similaridade consideravelmente alta com a seqüência de vídeo é possível, através de transformações geométricas, reconstruir o fundo de uma cena sem a necessidade da recomposição com o quadro original [6]. Assim seria possível montar o fundo dos quadros de um vídeo somente com a transmissão inicial do mosaico e com os parâmetros de movimento global no decorrer da transmissão. Diminuindo consideravelmente a quantidade de dados armazenados para a execução de um vídeo. Porém, é necessário que o mosaico seja muito bem estimado e que exista uma segmentação prévia dos objetos que estão na frente do fundo.

Um dos maiores problemas do mosaico é a sua construção [5]. Como foi mencionada anteriormente, a construção do mosaico possui duas etapas importantes:

- Estimação dos parâmetros de movimento global.
- Composição do mosaico, fusão dos quadros compensados.

Nas duas etapas, erros numéricos e erros de estimação são críticos na qualidade final do mosaico.

A etapa de estimação dos parâmetros de movimento global é um ponto crítico da geração do mosaico. Os parâmetros de movimento global podem apresentar erros, dependendo do método, devido à concatenação dos parâmetros de movimento quadro a quadro. Os erros devido à concatenação dos parâmetros

de movimento global podem, por muitas vezes, ser minimizados a partir de estimações feitas diretamente com o mosaico.

Um outro problema da estimação do movimento global baseada nos vetores de movimento é a confiabilidade destes. Os vetores de movimento dos blocos utilizados nas primeiras etapas de vários processos [5],[7],[8] propagam erros para as outras etapas da estimação de movimento global. Pois os processos clássicos de estimação local de movimento levam em consideração somente os valores de minimização ou maximizam valores locais, não considerando muitas vezes as regiões vizinhas. Para estimações de movimento global, mais importante do que o movimento de um bloco, de um quadro em relação a outro quadro, é o contexto em que o vetor de movimento está contido. Como estes processos não consideram o contexto em que o bloco está inserido, muitas vezes, geram erros na direção real do fluxo dos *pixels*. Alguns processos de estimação de movimento global tentam minimizar os efeitos da estimação local de movimentos, utilizando o campo de vetores de movimento somente como um ponto inicial da estimação [5],[7].

Na etapa de composição do mosaico temos outros problemas que causam erros na reconstrução dos quadros. Além dos problemas ocorridos devido à má estimação dos parâmetros de movimento global, temos o problema do encobrimento do fundo pelos objetos. Para resolver este problema é feita, em alguns processos [5], uma segmentação dos objetos que encobrem o fundo [3]. Como a segmentação, para objetos quaisquer, não é uma tarefa muito simples sempre existirão regiões de objetos compondo o fundo, acarretando erros na reconstrução do fundo da seqüência de vídeo.

Algumas vezes, no decorrer da seqüência de vídeo, ocorrem mudanças na iluminação. A mudança de iluminação é um problema sério na construção do mosaico. Com a mudança de iluminação a estimação de parâmetros é comprometida, pois os vetores de movimento não apontam para a direção real do movimento de câmera [9]. Mas existem alguns algoritmos que tentam minimizar os efeitos da diferença de iluminação [10].

O mosaico muitas vezes pode ter em sua construção erros devido à interpolação [11]. Como um *pixel* do mosaico, em alguns casos, é a composição de vários *pixels*, muitas vezes com níveis de cinza diferentes, no momento da reconstrução não é possível fazer o processo inverso sem perdas na qualidade do *pixel*. Portanto, mesmo com uma estimação de movimento global perfeita, o processo de interpolação provocará erros na construção do mosaico e na reconstrução do quadro a partir do mosaico e dos parâmetros de movimento. Uma forma de minimizar o erro devido à interpolação é utilizar algoritmos mais potentes de interpolação, que possuem complexidade computacional maior.

Outro grande problema da construção do mosaico é o tempo de processamento e a complexidade computacional [12]. Como a geração de mosaico necessita de processos matemáticos não muito triviais, o tempo de processamento é elevado. Alguns processos de geração de mosaico necessitam de busca completa [9], gradiente descendente [7] e outros processos custosos computacionalmente, na estimação do movimento global. Além dos processos de estimação de movimento global, as transformações geométricas também prejudicam a composição do mosaico nos quesitos de complexidade computacional e de tempo de processamento.

O método proposto neste trabalho é um novo esquema de geração do mosaico do fundo de uma seqüência de vídeo. O esquema é baseado na estimação do movimento global de cada quadro em relação ao mosaico. Sendo que, a estimação do movimento global é obtida através dos vetores de movimento locais dos blocos e da filtragem, retirada, dos vetores espúrios ou de vetores pouco confiáveis. Após a estimação dos parâmetros de movimento global é realizada a composição do mosaico. A composição do mosaico é realizada através da compensação do quadro, através dos parâmetros de movimento global estimados, seguido de uma média aritmética ponderada.

Este trabalho tem três objetivos:

- Primeiro, desenvolver um método de estimação de parâmetros de movimento global para aplicar na construção de imagens panorâmicas de fundo de uma seqüência de vídeo.

- Segundo, verificar a eficácia da estimação dos parâmetros de movimento global, para a obtenção de mosaicos.
- Terceiro, comparar com outros métodos já existentes, em termos da qualidade dos mosaicos estimados.

O trabalho está organizado da seguinte forma. No segundo capítulo são apresentados alguns trabalhos na área de estimação de movimento global e estimação de mosaico. O terceiro capítulo mostra a metodologia de estimação do mosaico. O quarto capítulo apresenta os resultados da utilização do método proposto. E no quinto e último capítulo é feita a conclusão do trabalho.

## 2. Métodos Aplicados à Estimação de Mosaico

Nesta seção serão apresentados alguns trabalhos na área de estimação de movimento global e na área de estimação de mosaico, para seqüências de vídeo. Existem vários autores, que apresentam estudos na área de extração de fundo de uma seqüência de vídeo e estimação de movimento global. Para a estimação de movimento global, alguns autores apresentam métodos baseados em fluxo ótico, outros com métodos baseados em gradiente descendente, além de outros métodos. Na questão de extração do fundo da seqüência de vídeo, existe uma variedade de métodos, que diferem desde a forma de compor o mosaico, até a forma de estimar o movimento global.

F. Dufaux e J. Konrad [7] propõem um método de estimação global de movimento automático, que utiliza um algoritmo baseado em uma pirâmide de imagens e gradiente descendente. A pirâmide é composta de dois quadros subseqüentes, com três resoluções diferentes cada quadro. O gradiente descendente é montado a partir das derivadas parciais dos erros da estimação do quadro em relação ao quadro real, pela variação de cada parâmetro de transformação. O processo inicia com a estimação dos valores da translação, por BMA (algoritmo por procura de bloco) em três passos, com os quadros de menor resolução. Após a estimação da translação, é utilizado um algoritmo de gradiente descendente, com os quadros de menor resolução, para a estimação e o refinamento de todos os parâmetros de movimento global. Os parâmetros de movimento global que foram encontrados são re-escalados, para a nova resolução dos quadros da etapa seguinte. Nas outras duas etapas o processo é o mesmo, exceto pelo BMA, com gradiente descendente seguido de re-escalamento dos parâmetros de movimento global. Outra proposta importante deste artigo é referente à estimação do erro, no gradiente descendente. Os erros, *pixel a pixel*, acima de um limiar não são considerados, pois podem fazer parte do movimento de objetos e não do fundo da imagem.

H. Cheng e W. Siu [13] propõem um método de estimação de movimento global baseado no artigo de Dufaux, citado acima, para geração de mosaico. Uma diferença entre o trabalho de Cheng e o de Dufaux é a presença do mosaico no processo de estimação do movimento global, na última etapa do processo. Outro diferencial deste método é a filtragem, retirada da estimação de movimento global, de regiões homogêneas e de regiões com muita textura. O restante do processo é muito semelhante ao apresentado por Dufaux.

C. R. Wren [14] propõe um método de estimação dos parâmetros de movimento global com o auxílio do usuário, que encontrará pontos similares em diferentes figuras. Este processo não é aplicado especificamente para vídeo. O algoritmo de estimação dos parâmetros movimento é baseado na solução de um sistema de equações, composto do deslocamento de cada ponto, escolhido pelo usuário em função dos parâmetros de movimento global. Para a resolução do sistema de equações é utilizado o método SVD (decomposição em valor singular). O artigo não propõe nenhum método de estimação de movimento por procura de bloco, para formar o sistema de equações. Porém tem como vantagem a solução do sistema de equações de uma forma mais simples e computacionalmente menos penosa que o gradiente descendente, por exemplo.

Outro trabalho importante, na parte de estimação dos parâmetros de movimento global, é o de F. Zhu, P. Xue e E. Ong [15]. Neste trabalho é utilizado um método de extração dos parâmetros de movimento global, através dos vetores de movimento. Neste caso, os vetores são obtidos de vídeos codificados pela transformada discreta dos co-senos, com compensação de movimento (MC-DCT). Com a ajuda dos coeficientes da DCT (transformada discreta de co-senos), são extraídos os vetores de movimento espúrios, sem uma reconstrução explícita da imagem. A filtragem dos vetores espúrios é feita através dos coeficientes DC da MC-DCT. Os valores DC maiores que o limiar são desconsiderados da estimação de movimento global. Os vetores restantes, os não espúrios, são usados no cálculo dos parâmetros de movimento global, através de mínimos quadrados.

Mais um trabalho importante em estimação de movimento global é apresentado por E. T. Kim e H. Kim [16]. Eles propõem um método de estimação

dos parâmetros, onde é realizada inicialmente uma seleção dos blocos. Os blocos selecionados são os que apresentam bordas ou altas texturas. Com os blocos selecionados, é feita a procura entre dois quadros consecutivos. Através dos deslocamentos estimados pela procura de blocos, são calculados os parâmetros de movimento, por mínimos quadrados. Os parâmetros de movimento global são utilizados na etapa de filtragem. Na filtragem são retirados os valores de deslocamento que estão excessivamente fora da tendência central estimada. Após a etapa de filtragem, é feita uma nova estimação dos parâmetros de movimento global. O método fica neste *loop* de cálculo dos parâmetros de movimento global e filtragem enquanto houver alguma modificação significativa.

G. Giunta e U. Mascia [17] mostram outra forma de estimação de movimento global, baseada em fluxo ótico. O primeiro passo do método é a estimação dos vetores de movimento, através de BMA (algoritmo de procura por bloco), com função de avaliação MAE (erro médio absoluto). O propósito do primeiro passo é minimizar os MAE's entre os blocos do quadro de referência e o quadro posterior. Encontrados os deslocamentos horizontais e verticais são montadas as equações de movimento global, com modelo parabólico, para cada bloco. Para avaliar a qualidade de cada equação de movimento são calculados os pesos, de cada expressão. Os pesos levam em conta a variação do erro médio absoluto pela variação do deslocamento dos blocos. Com os pesos de cada expressão calculados e as expressões de movimento dos blocos, são estimados os parâmetros de movimento global, através da solução do sistema de equações.

D. Wang e L. Wang [18] apresentam uma forma de estimação de movimento global semelhante à de outros trabalhos. Inicialmente é feita uma procura por similaridade de bloco, para estimar o campo de vetores de movimento. Do campo de vetores de movimento são calculadas as derivadas parciais espaciais dos vetores de movimento. Com as derivadas parciais espaciais de cada vetor calculadas são feitas as filtrações. As filtrações equivalem à retirada de algumas expressões, da avaliação do movimento global. O critério de retirada dos vetores de movimento é referente à média das derivadas parciais espaciais dos vetores de movimento. Caso o módulo da diferença entre a derivada parcial

espacial do vetor de movimento e o valor médio das derivadas parciais espaciais for maior do que o limiar o vetor não é considerado na estimação de movimento. Uma particularidade deste método é a utilização do valor médio das derivadas parciais espaciais dos vetores de movimento, após a filtragem, para o cálculo dos parâmetros de movimento.

Outra forma de estimar os parâmetros de movimento é apresentada por T.Chen e W. Lin e C. Chen [19], onde são utilizadas redes neurais para estimar o movimento global. O processo proposto por eles possui duas etapas. A primeira etapa consiste em uma procura por *pixels* semelhantes, entre dois quadros consecutivos, utilizando uma rede neural do tipo Hopfield. As redes neurais Hopfield são redes realimentadas, com treinamento não supervisionado e tem como objetivo minimizar uma função de energia. A função de energia tenta retratar numericamente a similaridade entre os blocos de *pixels* testados. Na segunda etapa, com os deslocamentos calculados na primeira etapa, são estimados os parâmetros de movimento. O cálculo dos parâmetros é feito através de uma rede neural *feed-forward* de três camadas, com treinamento supervisionado. O treinamento é realizado com seqüências de vídeo artificiais, em que os parâmetros de movimento já são conhecidos. Este artigo não apresenta especificamente um método para estimação de movimento global, movimento do fundo da seqüência de vídeo, mas sim um método de estimação de movimento rígido qualquer em três dimensões.

Em termos de estimação de mosaico o trabalho de Y. Lu, W. Gao e F. Wu [5] é um dos mais importantes. Neste trabalho é proposta uma codificação de vídeo com geração de *sprite* estático e também técnicas de codificação para formas arbitrárias. No método de geração de *sprite* estático, é realizada inicialmente uma segmentação dos objetos que estão à frente do fundo. A segmentação é feita através de máscaras auxiliares dos objetos, quando as máscaras estão disponíveis, ou através da segmentação de Rough, quando as máscaras não estão disponíveis. A segmentação de Rough é feita da seguinte forma. Inicialmente é realizada uma filtragem morfológica, na imagem de diferença entre o quadro de entrada e o mosaico transformado. No resultado da

filtragem é aplicado o operador de Canny [20], para uma extrair as bordas. E finalmente com o resultado do operador de Canny, é feito um preenchimento do interior das bordas, gerando dessa forma a máscara para segmentação. Com a máscara de segmentação pronta, são realizadas as classificações dos *pixels* do quadro de entrada em regiões realizáveis, regiões não realizáveis e regiões indefinidas. As regiões realizáveis são as regiões que fazem parte do fundo do quadro. As regiões não realizáveis são as regiões que não fazem parte do fundo do quadro. E as regiões indefinidas são as regiões de transição entre as regiões não realizáveis e as regiões realizáveis. Outra parte importante deste método de construção do mosaico de fundo é a estimação global de movimento. O estimador de movimento global deste artigo é baseado no método de Dufaux [7]. Porém com uma diferença, o método proposto neste trabalho, ao contrário ao de Dufaux, utiliza o mosaico como referência na estimação de movimento global. Segundo o artigo, a utilização do mosaico como referência no processo de estimação de movimento global, resolve o problema de acúmulo de erros da concatenação da estimação de movimento global quadro a quadro. Com os parâmetros de movimento global estimados, vem a fase da composição do mosaico. A composição do mosaico utiliza os valores dos parâmetros de movimento global estimados, junto com as máscaras de segmentação para a atualização do mosaico. As máscaras servem para distinguir a forma que cada *pixel* irá contribuir na composição do mosaico. Os *pixels* indefinidos não contribuem para a atualização do mosaico, os *pixels* realizáveis contribuem através de uma média ponderada e os *pixels* não realizáveis não são utilizados para compor o mosaico. A máscara é aplicada ao quadro, que é transformado, através dos parâmetros de movimento global, para compor o mosaico.

Outro trabalho importante para a estimação de mosaico é apresentado por F. Dufaux e F. Moscheni [21]. Onde os autores propõem uma nova técnica de construção do mosaico do fundo de uma seqüência de vídeo. O método é dividido em cinco partes estimação do movimento local, segmentação do fundo e do objeto, estimação de movimento global, compensação de movimento e composição do mosaico. A primeira parte do método, a estimação do movimento

local, é feita através de BMA busca completa com blocos de 8x8. A segunda parte é a diferenciação entre o fundo e os objetos, através das informações de movimento. A segmentação pelas informações de movimento utiliza o movimento local para diferenciar fundo dos objetos. Primeiramente é encontrado o movimento dominante de um conjunto de vetores de movimento, com o movimento dominante é feita uma filtragem dos vetores de movimento. Os vetores de movimento que seguem o movimento dominante são considerados como fundo e os que não seguem são considerados como objetos. Porém para fazer esta análise, é necessário considerar a falta de confiabilidade na direção dos vetores de movimento, em regiões com baixo contraste. A terceira parte deste método apresenta a estimação do movimento global, baseada em um trabalho de Moscheni, Dufaux e M. Kunt [14]. Esta estimação de movimento global é progressiva, onde primeiro são estimados os parâmetros de translação. Sendo os parâmetros de translação estimados através de BMA. Segundo são estimados os parâmetros de zoom, rotação e depois os parâmetros de perspectiva. O método também é multi-resolução, com a resolução dos quadros crescendo com o refinamento dos parâmetros. Os dois últimos estágios do método são a compensação do quadro com os parâmetros estimados no estágio anterior e a composição no mosaico. A composição é semelhante a uma média ponderada, porém não consideram na composição as regiões segmentadas inicialmente como objetos, através da segmentação pelos vetores de movimento.

A. Smolic, T. Sikora e J. Ohm [8] apresentam um método de estimação do mosaico, com uma forma diferente de estimação global de movimento. O método de estimação global deste artigo possui três etapas. Cada etapa tem como função refinar os parâmetros de movimento e aumentar gradativamente a complexidade do modelo de movimento global. Na primeira etapa é estimada a translação entre o quadro atual e o quadro estimado do mosaico, através de BMA. Sendo que o processo de BMA, neste artigo, é utilizado com um critério diferenciado de procura. Neste artigo é utilizada uma hessiana que, segundo os autores, é muito menos sensível às estimações espúrias do que o MSE (erro médio quadrático) e o MAE. Após a estimação da translação vem a segunda etapa, onde são estimados

os parâmetros de zoom, rotação e os parâmetros de translação são refinados. O processo da etapa dois é iterativo e baseado no fluxo ótico, campo de vetores de movimento. O primeiro passo desta etapa é a reconstrução do quadro atual, a partir do mosaico e dos parâmetros de movimento estimados, somente translação, na primeira etapa. Com o quadro reconstruído são estimados novos vetores de movimento em relação ao quadro atual. Com os novos vetores de movimento são estimados os parâmetros de movimento global, através de um sistema de equações. Este processo da etapa dois fica em *loop*, enquanto a diferença entre o quadro reconstruído e o quadro atual for maior que um limiar estabelecido. A etapa três é semelhante à etapa dois, porém com o diferencial do modelo de movimento utilizado. Na etapa três é utilizado o modelo de movimento global parabólico, que leva em conta transformações de mais alta ordem. Outra contribuição deste artigo é a classificação das regiões dos quadros do vídeo. Regiões em que somente existe o objeto são consideradas como não realizáveis. Regiões em que só existe a presença do fundo dos quadros são consideradas realizáveis e as regiões de transição entre o fundo e o objeto são consideradas indefinidas. A classificação das regiões dos quadros é utilizada no momento da composição do mosaico.

H. Cheung e W. Siu [12] apresentam um estimador de movimento global para a construção de imagens panorâmicas de fundo de uma seqüência de vídeo. O estimador de movimento proposto neste artigo tem como inovação a utilização de diversas referências para a estimação de movimento global. Em métodos anteriores é sempre utilizado o mosaico ou os parâmetros são concatenados quadro a quadro. Neste caso, o quadro de referência é modificado em duas situações. A primeira situação é quando a diferença, quadros visualmente diferentes, entre o quadro de teste e quadro de referência for muito grande. A outra situação de mudança do quadro de referência, é quando a diferença, em número de quadros, entre o quadro de teste e o quadro de referência for muito grande. Para o cálculo dos parâmetros de movimento global é utilizada procura por bloco seguida de gradiente descendente. Com os parâmetros de movimento global dos quadros calculados é feita a composição do mosaico. Para compor o

mosaico, é necessário que todas as estimações estejam na mesma referência. Para isso é feita uma concatenação dos parâmetros de movimento global do quadro de referência, em relação ao quadro inicial, com os parâmetros de movimento global do quadro de teste. Após a concatenação dos parâmetros, é feita a composição do mosaico através de média ponderada.

Neste capítulo foram apresentados alguns trabalhos na área de estimação de movimento global e na área de estimação de mosaico, para seqüências de vídeo. Os vários autores apresentam estudos na área de extração de fundo de uma seqüência de vídeo e estimação de movimento global. Para a estimação de movimento global os métodos baseados em fluxo ótico e métodos baseados em gradiente descendente são os mais usados. Na questão de extração do fundo da seqüência de vídeo existe uma variedade de métodos, que são diferentes na forma de estimação de movimento global e também na forma de compor o mosaico.

No próximo capítulo será mostrado o método de estimação de mosaico proposto neste trabalho. Sendo abordadas as técnicas de composição do mosaico e de estimação do movimento global.

## 3. Metodologia da estimação de mosaico

Nesta seção serão apresentados os métodos propostos que foram utilizados neste trabalho. Primeiro será apresentada uma visão geral do método de estimação de movimento global, como é computado e as funções de cada parte do método. Na segunda parte serão apresentadas as especificações do método de estimação de movimento global. Na terceira parte será apresentado o método de composição do mosaico. E na quarta parte serão apresentados os padrões e critérios de desempenho utilizados.

### 3.1. Método de Estimação de Movimento Global

O método de estimação de movimento proposto neste trabalho tem como ponto inicial a estimação de movimento por procura de bloco, BMA. Com os vetores de movimento obtidos, é realizada uma filtragem de possíveis vetores de movimento espúrios. O campo de vetores estimado, pela filtragem, é utilizado para a estimação de movimento global.

Este estimador de movimento global foi criado com o objetivo de atenuar a interferência de vetores de movimento espúrios no resultado final da estimação. São considerados vetores espúrios os vetores que fogem da tendência central do movimento global. Também são considerados vetores espúrios os vetores de regiões muito homogêneas, em que a tomada de decisão de qual o vetor minimiza o MSE não é muito clara.

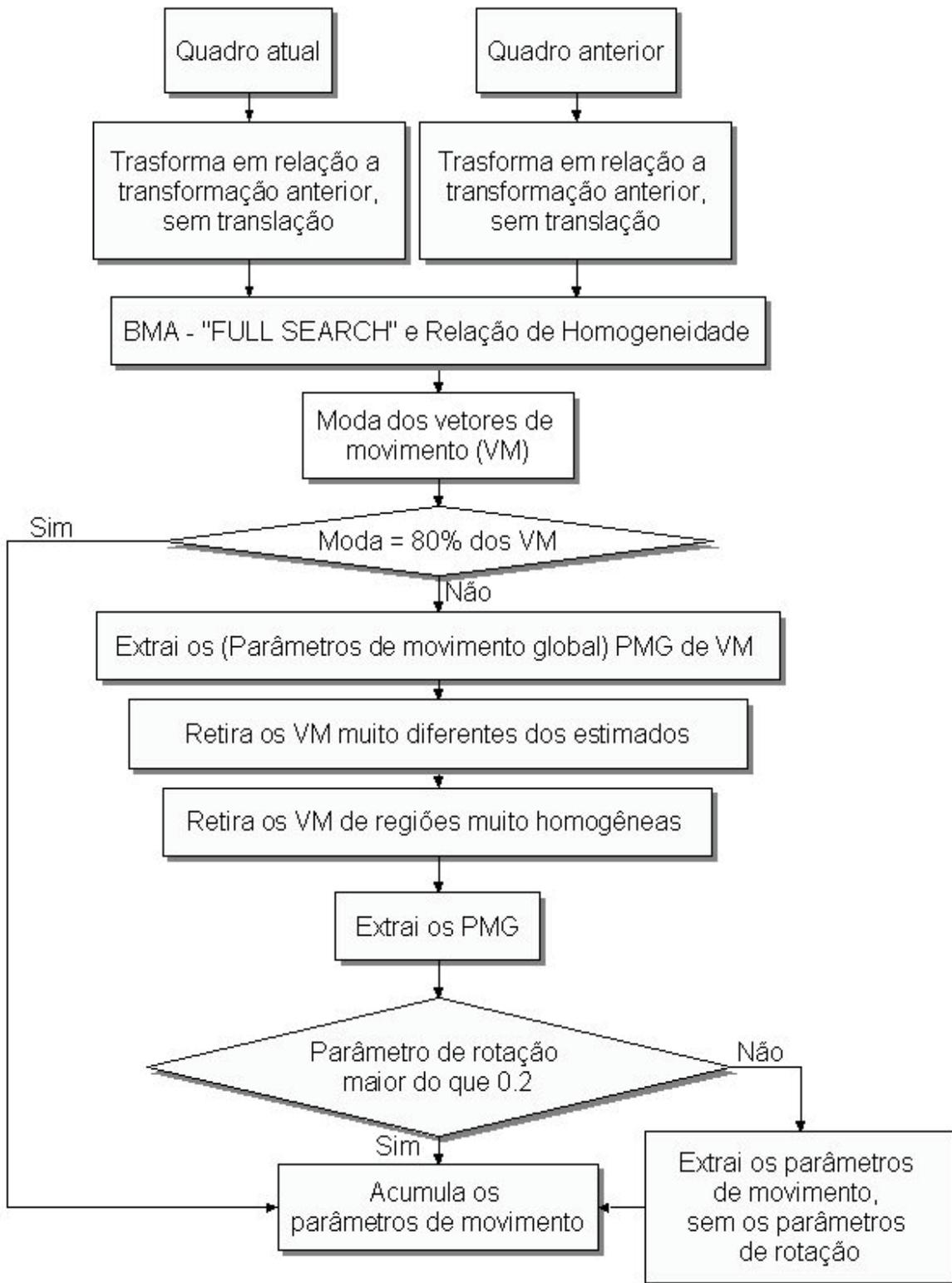
Os vetores considerados espúrios são filtrados, retirados, das etapas subseqüentes de estimação de movimento global. Esses vetores são retirados por não possuírem uma informação confiável da real direção do fluxo dos *pixels* do fundo da seqüência de vídeo. Os vetores de movimento que fogem muito da tendência central do campo de vetores podem, muitas vezes, representar o fluxo de *pixels* de objetos que se movem de forma diferente do movimento global, movimento do fundo do quadro. Vetores de regiões muito homogêneas não

retratam, com exatidão, a direção do vetor de movimento. A influência dos vetores mal estimados, dependendo do vídeo, na estimação do movimento global é significativa. Sendo necessária a correção da direção dos vetores ou a retirada de alguns vetores de movimento. A retirada de alguns dos vetores de movimento é possível, graças à redundância presente no campo de vetores. Não sendo necessária a presença de todos os vetores de movimento para estimar os parâmetros de movimento global.

Este processo de estimação de movimento global é realizado uma única vez, pois esta baseado em uma de seleção dos vetores de movimento que exige algumas características muito rígidas, tendo poucas melhoras com outras iterações.

O algoritmo de estimação tem um comportamento eficiente, porém no quesito tempo de processamento deixa um pouco a desejar. O grande problema do processo é a necessidade de ter um campo de vetores com uma qualidade razoável. Sendo que a qualidade dos campos de vetores está intimamente relacionada com o método de procura por blocos inicial. Métodos mais rápidos de procura de blocos, como o método de três passos, por exemplo, em alguns casos, não possuem campos coesos o bastante para uma boa estimação de movimento global. Devido à necessidade de campos coesos foi optado por processos com menor desempenho em termos de tempo de processamento. Porém com um comprometimento maior com a qualidade dos vetores de movimento.

O método de estimação dos parâmetros de movimento global é mostrado pelo fluxograma da Figura 3:



**Figura 3-** Fluxograma do método de estimação de movimento global.

A estimação dos parâmetros de movimento global é feita do seguinte modo:  
-Primeiro, o quadro atual e o anterior são transformados pelos parâmetros de movimento global estimados para o quadro anterior, sem considerar a translação, com os parâmetros de translação nulos. As duas imagens são transformadas com o propósito de não propagar erros devido ao acúmulo dos parâmetros de movimento.

-Segundo, são estimados os vetores de movimento e a relação de homogeneidade das regiões de procura entre as duas imagens consecutivas. Os vetores de movimento são estimados através de busca completa. A escolha do método busca completa é devido a melhor qualidade dos vetores de movimento estimados, comparados com a procura em três passos.

A relação de homogeneidade é calculada pela diferença entre os valores de erro médio quadrático (MSE), dos diversos possíveis vetores, na região de procura. Caso a diferença entre os MSE devido ao vetor escolhido e devido a outro vetor, dentro da região de procura seja pequena, o vetor escolhido é considerado impróprio para a extração de parâmetros em futuros cálculos.

-Terceiro, é calculada a moda dos vetores de movimento. Caso a moda corresponda a mais de 80%, do número total de vetores considerando até os vetores excluídos, o movimento que a moda representa é considerado como movimento global. Em caso contrário, o processo segue sem nenhuma contribuição da moda.

-Quarto, através de mínimos quadrados são extraídos os parâmetros de movimento global. Os parâmetros são estimados através de um sistema de equações que envolvem todos os vetores de movimento inclusive os vetores de regiões muito homogêneas.

-Quinto, utilizando os parâmetros de movimento global é criado um campo de vetores artificial. O campo de vetores estimado é comparado com o campo de vetores artificial, caso a diferença entre os dois vetores seja grande o vetor

estimado é retirado da próxima etapa da estimação. Este passo tem como objetivo retirar os vetores correspondentes aos objetos da estimação de movimento.

-Sexto, são estimados, através de mínimos quadrados, os parâmetros de movimento sem os vetores que fogem muito da tendência central e sem os vetores de regiões muito homogêneas. Neste passo são calculados os parâmetros de movimento em sua última aproximação, porém sem verificar se os valores têm algum erro numérico de estimação.

-Sétimo, estimados os novos parâmetros, é verificado se os parâmetros de rotação e de perspectiva são menores do que um certo limiar. Sendo menores que este certo limiar são extraídos do cálculo das novas etapas da estimação dos parâmetros. O limiar nesse caso é de 0.001 para os parâmetros de perspectiva e 0.01 para os parâmetros de rotação. O limiar é utilizado para evitar que os erros numéricos se propaguem.

-Oitavo, caso algum parâmetro seja irrelevante ele é extraído desta etapa de estimação e um novo cálculo dos parâmetros de movimento global, através de mínimos quadrados, é feito. Esta etapa tem como propósito evitar a propagação de erro devido a pequenos valores dos parâmetros de movimento.

Uma das limitações deste método é referente aos objetos grandes, na frente do fundo do quadro. Com objetos grandes, em relação ao tamanho do quadro, existirá uma quantidade grande de vetores de movimento referentes aos objetos, atrapalhando na distinção entre objetos e fundo. Pois o método faz distinção entre objetos e fundo do quadro através da tendência de movimento central de todo o conjunto. Dessa forma, o algoritmo irá confundir fundo com o objeto. Outro problema é o dos objetos que não se movem na mesma direção e sentido dos vetores do fundo, porém estão muito próximos do movimento fundo. Apesar de estarem próximos estes vetores geram erros de estimação, que acumulados podem tornar a estimação pouco precisa.

## 3.2. Pontos específicos do método de estimação do movimento global

### 3.2.1. Modelos de movimento global

Neste trabalho foi utilizado um modelo de movimento global com perspectiva [5], porém em casos especiais foram utilizadas formas simplificadas do modelo. O modelo com perspectiva tenta simular a rotação, zoom, translação e perspectiva. O modelo de movimento utilizado é regido pelas equações (1) e (2).

$$X = \frac{a_0 \cdot x + a_2 \cdot y + a_4}{a_6 \cdot x + a_7 \cdot y + 1} \quad (1)$$

$$Y = \frac{a_1 \cdot x + a_3 \cdot y + a_5}{a_6 \cdot x + a_7 \cdot y + 1} \quad (2)$$

Onde  $x$  e  $y$  são as coordenadas originais,  $X$  e  $Y$  são as coordenadas transformadas,  $a_4$  e  $a_5$  são parâmetros de translação,  $a_0$  e  $a_3$  são parâmetros de zoom,  $a_2$  e  $a_1$  parâmetros de rotação,  $a_6$  e  $a_7$  parâmetros de perspectiva.

O modelo de movimento global com oito parâmetros, para movimentos de câmera, é adequado para este trabalho. Porém, em situações de movimentos mais complexos, como distorções parabólicas, o modelo não possui uma modelagem muito fiel. Portanto, este modelo é mais utilizado para vídeos naturais, feitos com câmera ou para vídeos artificiais que simulem o movimento de uma câmera.

Além da questão da fidelidade com os movimento de uma câmera, existe a questão da sensibilidade dos parâmetros de movimento global. Os termos do modelo de movimento  $a_6$  e  $a_7$  possuem uma sensibilidade alta. Esta alta sensibilidade dos parâmetros que estão no denominador decorre dos parâmetros possuírem valores com uma ordem de grandeza pequena, se comparados com os

outros parâmetros do modelo, pois uma pequena variação desses parâmetros de movimento pode gerar transformações completamente diferentes. Portanto, em situações em que a estimação de movimento global não é muito fiel, esses parâmetros mais atrapalham do que auxiliam. Para estas situações foi optado utilizar um modelo mais simples com apenas seis parâmetros,

$$X = a_0 \cdot x + a_2 \cdot y + a_4, \quad (3)$$

$$Y = a_1 \cdot x + a_3 \cdot y + a_5, \quad (4)$$

objetivando uma melhor qualidade do mosaico.

O modelo de movimento global com seis parâmetros não consegue modelar perspectiva, porém é eficaz para translação, rotação e zoom. Portanto em situações em que os parâmetros de perspectiva não são muito exigidos podemos muito bem modelar o movimento global com apenas seis parâmetros.

Para situações mais simples, em que é percebido que o movimento é de translação pura, o modelo de movimento global é apresentado da seguinte forma,

$$X = 1 \cdot x + 0 \cdot y + a_4, \quad (5)$$

$$Y = 0 \cdot x + 1 \cdot y + a_5. \quad (6)$$

Onde são estimados somente os parâmetros de translação, sendo o restante da equação desconsiderada.

### **3.2.2. Transformação geométrica (*warping*)**

A transformação geométrica [22] é um ponto importante deste trabalho, pois é através dela que o quadro é compensado para ser fundido com o mosaico. Este processo de transformação geométrica também é utilizado para a estimação dos quadros a partir do mosaico, porém os papéis do mosaico e do quadro são invertidos.

Para realizar a transformação geométrica direta é encontrada primeiramente a expressão da transformação inversa,

$$x = \frac{-(a_3 \cdot X - a_7 \cdot a_5 \cdot X + a_2 \cdot a_5 + a_4 \cdot a_7 \cdot Y - a_4 \cdot a_3 - a_2 \cdot Y)}{X \cdot a_6 \cdot a_3 + a_7 \cdot a_0 \cdot Y - a_0 \cdot a_3 - a_2 \cdot a_6 \cdot Y - a_1 \cdot a_7 \cdot X + a_1 \cdot a_2}, \quad (7)$$

$$y = \frac{-(a_6 \cdot a_5 \cdot X - a_1 \cdot X - a_0 \cdot a_5 - a_4 \cdot a_6 \cdot Y + a_4 \cdot a_1 + a_0 \cdot Y)}{X \cdot a_6 \cdot a_3 + a_7 \cdot a_0 \cdot Y - a_0 \cdot a_3 - a_2 \cdot a_6 \cdot Y - a_1 \cdot a_7 \cdot X + a_1 \cdot a_2}. \quad (8)$$

As expressões (7) e (8) mostram a expressão inversa a partir da transformação direta estimada.

Com a expressão da transformada inversa encontramos os pontos extremos dentro do mosaico e é feito o mapeamento dos *pixels* do quadro no mosaico. Os pontos extremos são encontrados através dos pontos limites do quadro, pois as transformações são lineares. Encontrados os limites do mapeamento no mosaico, é feita a transformação inversa, do mosaico para o quadro, *pixel a pixel* dentro dos limites encontrados no mosaico. Em alguns casos, os pontos mapeados na imagem original não estão em coordenadas inteiras. Nestes casos utilizamos uma interpolação bilinear para encontrar os tons de cinza correspondentes para a composição do mosaico.

A interpolação bilinear é calculada do seguinte modo.

- Primeiro são encontrados os valores dos maiores inteiros menores que o valor da coordenada que será interpolada, truncamento do valor,

$$x' = \text{trunc}(x), \quad (9)$$

$$y' = \text{trunc}(y), \quad (10)$$

onde  $x'$  e  $y'$  são as coordenadas mapeadas na imagem original truncadas.

-Segundo são encontrados os parâmetros para o futuro cálculo dos pesos, esses parâmetros estão relacionados com a distância entre o ponto real e o maior inteiro menor que o ponto real,

$$t = x' - x \quad (11)$$

$$u = y' - y \quad (12)$$

-Terceiro é calculada a média, ponderada pelos parâmetros  $t$  e  $u$ , para encontrar o nível de cinza da coordenada interpolada,

$$I''(X,Y) = \begin{cases} (1-t).(1-u)I(x',y') + t.(1-u)I(x'+1,y') + \\ t.u.I(x'+1,y'+1) + (1-t).u.I(x',y'+1) \end{cases} \quad (13)$$

Onde  $I''(X,Y)$  e  $(X,Y)$  são o *pixel* do quadro interpolado e a coordenada do *pixel* do quadro interpolado, respectivamente. E  $I(x,y)$ ,  $I(x+1,y)$ ,  $I(x,y+1)$  e  $I(x+1,y+1)$  são os pontos que circundam as coordenadas mapeadas na imagem original  $(x,y)$ .

Os pixels, da região delimitada dentro do mosaico, que não são mapeados dentro dos limites do quadro, não são considerados na composição do mosaico.

A interpolação bilinear foi utilizada devido a sua computação mais simples do que outras formas de interpolação, como a interpolação bicúbica. Em termos de qualidade muitas vezes deixa a desejar, porém tem uma boa relação custo computacional por qualidade da imagem.

### 3.2.3. Algoritmo de Procura por bloco (BMA) busca completa

O algoritmo de procura por bloco utilizado neste trabalho é o busca completa [9]. O busca completa verifica todos os possíveis blocos de uma região de busca. No método busca completa o objetivo é encontrar o bloco que maximize ou minimize um certo critério dentro do campo de busca. Neste trabalho o método de busca completa é utilizado para encontrar o bloco que minimize o MSE [22]

$$MSE(a,b) = \frac{1}{MN} \sum_{x=x_1}^{x_1+M-1} \sum_{y=y_1}^{y_1+N-1} [I(x,y) - I'(x+a,y+b)]^2, \quad (14)$$

dentro do campo de busca. Onde  $I$  é o quadro atual,  $I'$  é o quadro anterior,  $M$  e  $N$  são as dimensões do bloco,  $a$  e  $b$  é a translação do bloco  $I'$  em relação a origem do bloco  $I$  e  $x_1$  e  $y_1$  é a origem do bloco do quadro  $I$ .

Comparado com outros métodos de procura de bloco o busca completa é bem mais lento, porém é o método que gera um campo de vetores mais coeso. Para este caso é interessante que os vetores, além de apontarem para o bloco mais semelhante, sigam a tendência do movimento dos outros vetores do campo. Pois o objetivo da extração do campo de vetores é para uma futura estimação dos parâmetros de movimento global.

Outros critérios de escolha de bloco, além do MSE, foram testados como o erro médio absoluto [22],

$$MAE(a,b) = \frac{1}{MN} \sum_{x=x_1}^{x_1+M-1} \sum_{y=y_1}^{y_1+N-1} |I(x,y) - I'(x+a,y+b)|, \quad (15)$$

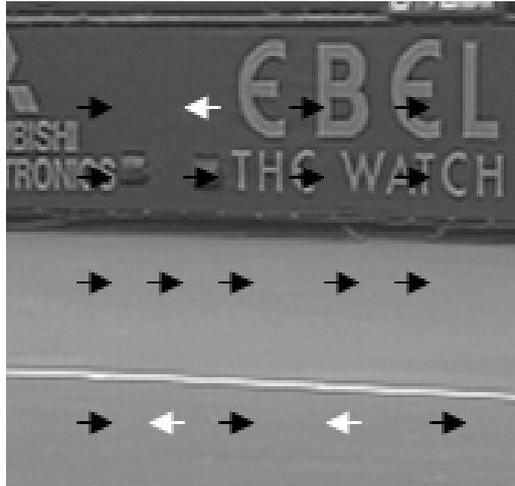
e a correlação cruzada CC [23],

$$CC(a,b) = \sum_{x=x_1}^{x_1+M-1} \sum_{y=y_1}^{y_1+N-1} |I(x,y) \cdot I'(x+a,y+b)|. \quad (16)$$

Porém, apresentaram desempenhos inferiores ao MSE. Outra razão da utilização do MSE é devido a sua complexidade computacional mais simples do que a correlação cruzada.

### 3.2.4. Relação de homogeneidade

O critério de homogeneidade foi criado devido à inconsistência na direção de alguns vetores de movimento. Em regiões de textura homogênea, alguns vetores de movimento têm comportamento muito diferente dos vetores de movimento vizinhos, mesmo sem objetos próximos. Este problema acontece em muitos trechos de várias seqüências. Um exemplo deste problema é mostrado na Figura 4, que é um trecho do quadro 10 da seqüência Stefan.



**Figura 4-** Vetores de movimento de regiões homogêneas.

Na Figura 4 é possível ver que alguns vetores estão apontando para o lado esquerdo, sendo que neste caso o movimento global era para a direita. Dessa forma, foi necessário criar um critério de qualidade da estimação, relacionado com a homogeneidade da superfície de MSE's da região de procura. Outra questão fundamental foi encontrar até que ponto a pequena diferença entre os MSE's, da região de procura, prejudica a qualidade da estimação do vetor de movimento.

A primeira parte do cálculo da relação de homogeneidade é o calculo da superfície de MSE's, SMSE,

$$SMSE = \begin{bmatrix} MSE_{00} & \dots & MSE_{0n} \\ : & \dots & : \\ MSE_{n0} & \dots & MSE_{nm} \end{bmatrix}. \quad (17)$$

A SMSE representa todos os valores de MSE's dentro da região de procura, macrobloco, onde  $n$  representa o tamanho da janela de procura. Após a criação da SMSE, é calculada uma medida de dispersão  $D$ ,

$$D = \begin{bmatrix} D_{00} & \dots & D_{0n} \\ : & \dots & : \\ D_{n0} & \dots & D_{nm} \end{bmatrix}, \quad (18)$$

onde cada termo da matriz de dispersão é calculada a partir da expressão:

$$D(i, j) = \frac{|\min(SMSE) - MSE(i, j)|}{\min(SMSE)}. \quad (19)$$

Onde  $\min(SMSE)$  corresponde ao menor valor de  $MSE$  de  $SMSE$  e  $MSE(i, j)$  é o termo da posição  $(i, j)$  da matriz  $SMSE$ .

A intenção é verificar se o vetor escolhido não possui um outro vetor possível, vetor concorrente, dentro da região de procura, com  $MSE$  muito próximo. A intenção é retirar os vetores, das próximas etapas, em que a escolha não é muito clara. Para verificar se a matriz de dispersão possui valores que tornam a medida de direção e sentido do vetor confiável, é feita a *EBSLD* (escolha baseada na soma dos limiares das dispersões).

Para cada termo da matriz de dispersões é calculada uma verificação de limiar,

$$L(i, j) = \begin{cases} 0, & D(i, j) > \text{limiar} \\ 1, & D(i, j) \leq \text{limiar} \end{cases}, \quad (20)$$

com todas as verificações de limiar é montada outra matriz, a matriz das verificações de limiares,

$$L = \begin{bmatrix} L_{00} & \dots & L_{0n} \\ \vdots & \dots & \vdots \\ L_{n0} & \dots & L_{nn} \end{bmatrix}. \quad (21)$$

Com os limiares calculados é feita a soma *SLD* (soma dos limiares das dispersões):

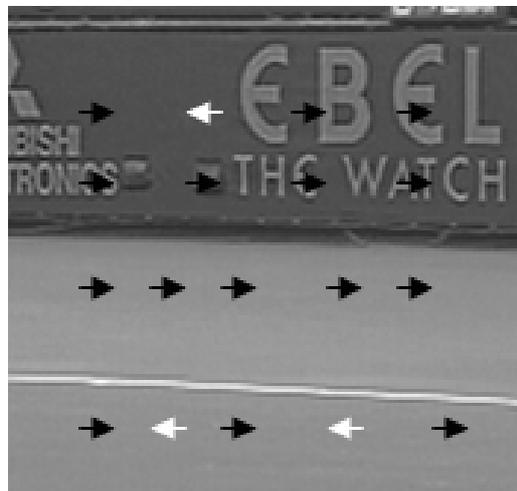
$$SLD = \sum_{i=0}^n \sum_{j=0}^n L(i, j) \quad (22)$$

Com o valor de SLD é feita a análise de confiabilidade do vetor de movimento, onde EBSLD pode ter dois valores, confiável,  $EBSLD = 0$ , ou não confiável,  $EBSLD = 1$ , através da expressão:

$$EBSLD(SLD) = \begin{cases} \text{confiável,} & SLD = 1 \\ \text{não confiável,} & SLD > 1 \end{cases} \quad (23)$$

Experimentalmente foi verificado que os valores de dispersões percentuais, entre o vetor escolhido e os outros possíveis vetores de movimento, menores que 0,02 são relevantes para a confiabilidade da estimação. A confiabilidade foi verificada através da qualidade visual do mosaico. Com esta verificação foi atribuído o valor de 0,02 de dispersão como valor de corte, para excluir os vetores de movimento da estimação do movimento global.

A ação do método de classificação pode ser observada na Figura 5.



**Figura 5-** Classificação dos vetores de movimento.

Os vetores claros são os vetores impróprios para a estimação e os vetores escuros são aptos para a estimação do movimento global.

Além da homogeneidade, existem outros problemas que dificultam na escolha correta da direção do vetor de movimento e que não são tratados neste

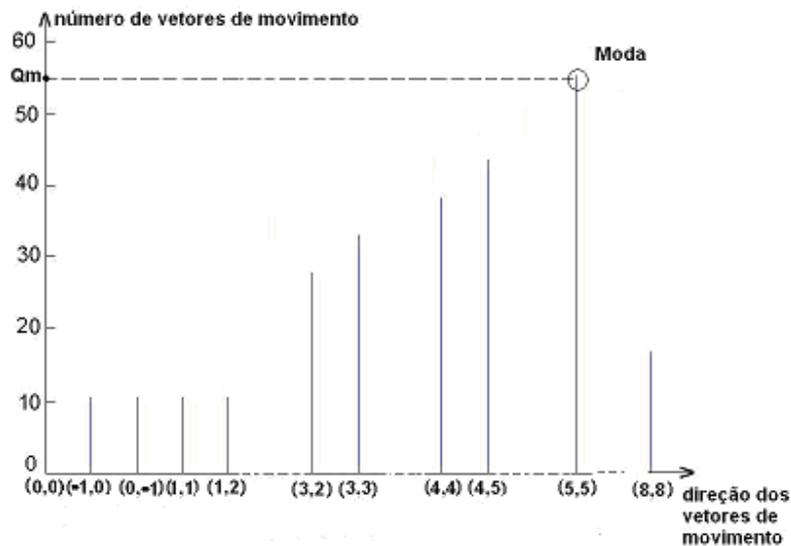
método. Entre os problemas que não são tratados neste método é possível citar, a diferença de iluminação entre quadros consecutivos e também o pequeno número de níveis de quantização na captura do vídeo.

### 3.2.5. Translação pura

O método possui uma forma de estimar translação pura. A estimacão é baseada na quantidade percentual de vetores do campo que estão presentes na moda do campo de vetores de movimento  $CV$ ,  $QModa$ ,

$$QModa = \frac{Qm(moda(CV))}{Ncv} \cdot 100\% . \quad (24)$$

Onde  $Qm$  retorna o número de vetores presentes na moda de  $CV$ , como mostra a Figura 6 e  $Ncv$  é o número de vetores do campo  $CV$ .



**Figura 6-** Exemplo de histograma, para a obtenção da moda

O valor de  $QModa$  é comparado com um limiar, experimentalmente encontrado através da função  $Tr$ , onde

$$Tr(QModa) = \begin{cases} 1, & QModa > \text{limiar} \\ 0, & QModa \leq \text{limiar} \end{cases}, \quad (25)$$

caso a moda for superior ao limiar, ela será considerada o movimento dominante. E os parâmetros de movimento serão estimados somente em função dos vetores de movimento que estão presentes na moda. Caso contrário esta etapa não terá influência no restante das etapas de estimação de movimento global.

### 3.2.6. Montagem do sistema de equações

Para a montagem do sistema de equações, que encontra os parâmetros de movimento global, primeiramente são manipuladas as expressões (7) e (8) transformadas em:

$$X = x \cdot a_0 + y \cdot a_2 + 1 \cdot a_4 + 0 \cdot a_1 + 0 \cdot a_3 + 0 \cdot a_5 - X \cdot x \cdot a_6 - X \cdot y \cdot a_7 \quad (26)$$

$$Y = 0 \cdot a_0 + 0 \cdot a_2 + 0 \cdot a_4 + x \cdot a_1 + y \cdot a_3 + 1 \cdot a_5 - Y \cdot x \cdot a_6 - Y \cdot y \cdot a_7. \quad (27)$$

Onde X e Y representam as coordenadas da posição que o vetor de movimento aponta no quadro atual e x e y são as coordenadas da origem do vetor que estão no quadro anterior.

Cada vetor de movimento gera duas expressões como as equações (26) e (27). Com todas as expressões geradas é montado um sistema de equações, baseado no trabalho de E. T. Kim [6], como:

$$\begin{bmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 & -x_1 \cdot X_1 & -y_1 \cdot X_1 \\ 0 & 0 & 0 & x_1 & y_1 & 1 & -x_1 \cdot Y_1 & -y_1 \cdot Y_1 \\ x_2 & y_2 & 1 & 0 & 0 & 0 & -x_2 \cdot X_2 & -y_2 \cdot X_2 \\ 0 & 0 & 0 & x_2 & y_2 & 1 & -x_2 \cdot Y_2 & -y_2 \cdot Y_2 \\ \vdots & & & \vdots & & & \vdots & \\ \vdots & & & \vdots & & & \vdots & \\ x_n & y_n & 1 & 0 & 0 & 0 & -x_n \cdot X_n & -y_n \cdot X_n \\ 0 & 0 & 0 & x_n & y_n & 1 & -x_n \cdot Y_n & -y_n \cdot Y_n \end{bmatrix} \cdot \begin{bmatrix} a_0 \\ a_2 \\ a_4 \\ a_1 \\ a_3 \\ a_5 \\ a_6 \\ a_7 \end{bmatrix} = \begin{bmatrix} X_1 \\ Y_1 \\ X_2 \\ Y_2 \\ \vdots \\ \vdots \\ X_n \\ Y_n \end{bmatrix} \quad (28)$$

Em algumas etapas do processo de estimação algumas expressões são retiradas do sistema de equações. As situações em que ocorre a retirada de algumas equações são atribuídas pelos processos de classificação dos vetores.

### 3.2.7. Resolução dos sistemas de equações

Para resolver o sistema de equações foi utilizado o método SVD [24]. Este método tem como três principais atrativos. Primeiro, uma boa estabilidade numérica em relação à outros métodos. Segundo, a possibilidade de utilizar um número de equações muito superior ao número de variáveis. Terceiro, mesmo não existindo uma resposta ideal do sistema de equações é encontrada uma resposta aproximada que minimiza  $\varepsilon$  onde

$$\varepsilon = |H - \delta \cdot b|. \quad (29)$$

Sendo  $H$  a matriz que representa cada vetor de movimento,  $b$  os parâmetros de movimento global e  $\delta$  são os pontos futuros de deslocamento.

### 3.2.8. Retirada de vetores de movimento de objetos

Vetores de movimento de objetos interferem na estimação do movimento global [15], principalmente se o movimento do objeto não tiver um comportamento semelhante ao fundo. Para contornar este problema foi desenvolvido um método para retirar os vetores de objetos da estimação de movimento global, semelhante ao utilizado em outros métodos [13]. O método se baseia no comportamento geral do campo de vetores e na hipótese que a área que o objeto possui no quadro é bem menor que a área ocupada pelo fundo.

O método segue a seguinte ordem:

-Primeiro, do campo de vetores  $CV$  são extraídos os parâmetros de movimento global, através de mínimos quadrados. Onde  $CV$  são duas matrizes com as direções dos vetores de movimento em  $x$ ,  $CV_x$ , e  $y$ ,  $CV_y$ ,

$$CV_x = \begin{bmatrix} CV_{x_{00}} & \dots & CV_{x_{0n}} \\ \vdots & \dots & \vdots \\ CV_{x_{n0}} & \dots & CV_{x_{nn}} \end{bmatrix}, \quad (30)$$

$$CV_y = \begin{bmatrix} CV_{y_{00}} & \dots & CV_{y_{0n}} \\ \vdots & \dots & \vdots \\ CV_{y_{n0}} & \dots & CV_{y_{nn}} \end{bmatrix}, \quad (31)$$

onde cada termo das matrizes corresponde a direção de um vetor de movimento, sendo,

$$CV = (CV_x, CV_y). \quad (32)$$

-Segundo, a partir dos parâmetros de movimento é gerado um campo de vetores de movimento artificial ACV, com a mesma configuração de CV,

$$ACV_x = \begin{bmatrix} ACV_{x_{00}} & \dots & ACV_{x_{0n}} \\ \vdots & \dots & \vdots \\ ACV_{x_{n0}} & \dots & ACV_{x_{nn}} \end{bmatrix}, \quad (33)$$

$$ACV_y = \begin{bmatrix} ACV_{y_{00}} & \dots & ACV_{y_{0n}} \\ \vdots & \dots & \vdots \\ ACV_{y_{n0}} & \dots & ACV_{y_{nn}} \end{bmatrix}. \quad (34)$$

-Terceiro são comparados os termos de ACV com CV, através da aplicação da função:

$$VOB(i, j) = \begin{cases} 1, & \sqrt{(ACV_x(i, j) - VC_x(i, j))^2 + (ACV_y(i, j) - VC_y(i, j))^2} \geq \text{limiar} \\ 0, & \sqrt{(ACV_x(i, j) - VC_x(i, j))^2 + (ACV_y(i, j) - VC_y(i, j))^2} < \text{limiar} \end{cases}, \quad (35)$$

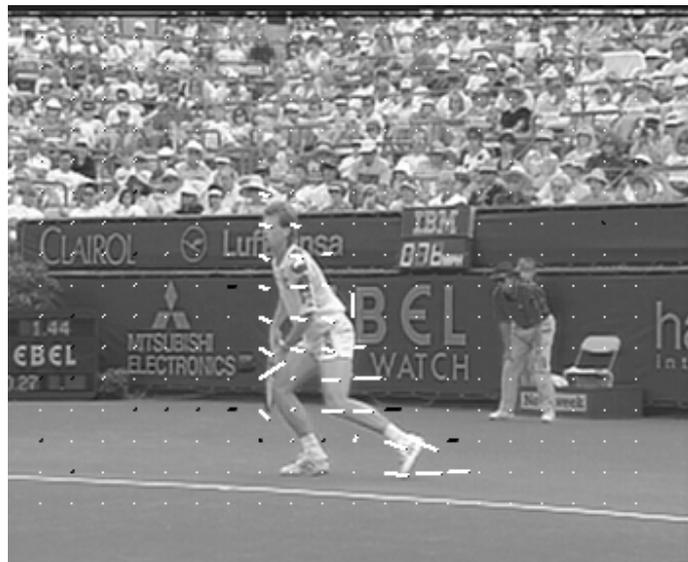
onde  $VOB(i, j)$  é uma função que indica se um vetor na posição  $(i, j)$  é de objeto ou não. Se a resposta de  $VOB(i, j)$  for 1, o vetor é de objeto, se a resposta for 0, o vetor não será de objeto. O VOB é calculado para cada vetor, gerando uma matriz

que identifica os vetores de movimento de objetos. Os vetores identificados com um, na matriz  $VOB$ , não serão utilizados na estimação global de movimento,

$$VOB = \begin{bmatrix} 1 & \dots & \dots & 0 \\ 0 & \dots & \dots & 0 \\ \vdots & \dots & \dots & \vdots \\ 0 & \dots & \dots & 1 \end{bmatrix}. \quad (36)$$

Como foi comentado acima é necessário que o objeto seja pequeno em comparação com o fundo. Pois o método está baseado na exclusão de vetores de movimento que não seguem a estimativa da tendência central de movimento global.

A Figura 7 exemplifica a ação do algoritmo na extração dos vetores de movimento de objetos.



**Figura 7-** Extração dos vetores de movimento dos objetos.

Na Figura 7 é possível ver os vetores de movimento do objeto, neste caso os vetores que acompanham o jogador Stefan, que são os vetores identificados pela cor branca próximos ao jogador.

### 3.2.9. Acúmulo de erros na estimação de movimentos

Quando utilizamos estimações de movimento global quadro a quadro, é necessário concatenar as transformações para encontrar a transformação do quadro  $n$  em relação ao primeiro quadro, como mostrado no trabalho de N. Grammalidis, D. Beletsiotis e M. Strintzis [23].

$$\begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} = \begin{bmatrix} a_0^0 & a_2^0 & a_4^0 \\ a_1^0 & a_3^0 & a_5^0 \\ 0 & 0 & 1 \end{bmatrix} \dots \dots \begin{bmatrix} a_0^{n-1} & a_2^{n-1} & a_4^{n-1} \\ a_1^{n-1} & a_3^{n-1} & a_5^{n-1} \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} a_0^n & a_2^n & a_4^n \\ a_1^n & a_3^n & a_5^n \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}, \quad (37)$$

onde a transformação equivalente seria,

$$\begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} = \begin{bmatrix} a_0^{eq} & a_2^{eq} & a_4^{eq} \\ a_1^{eq} & a_3^{eq} & a_5^{eq} \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}. \quad (38)$$

Porém devido à concatenação existe um acúmulo de erros na estimação dos parâmetros de movimento global [5]. Os parâmetros de movimento, em algumas situações, acumulam pequenos valores que não representam a transformação efetiva. Com o aumento do número de quadros o acúmulo desses erros é significativo, ocasionando problemas no valor do parâmetro e distorcendo muitas vezes o mosaico. Esta constatação pode ser demonstrada matematicamente, da seguinte maneira. Considerando um modelo de movimento mais simples com seis parâmetros de movimento, a concatenação dos parâmetros de movimento global, de dois quadros se comporta desta maneira:

$$\begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} = \begin{bmatrix} a_0^0 & a_2^0 & a_4^0 \\ a_1^0 & a_3^0 & a_5^0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} a_0^1 & a_2^1 & a_4^1 \\ a_1^1 & a_3^1 & a_5^1 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}, \quad (39)$$

encontrando a transformação equivalente,

$$\begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} = \begin{bmatrix} a_0^{eq} & a_2^{eq} & a_4^{eq} \\ a_1^{eq} & a_3^{eq} & a_5^{eq} \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}, \quad (40)$$

onde

$$a_0^{eq} = a_0^0 \cdot a_0^1 + a_2^0 \cdot a_1^1, \quad (41)$$

$$a_2^{eq} = a_0^0 \cdot a_2^1 + a_2^0 \cdot a_3^1, \quad (42)$$

caso  $a_0^0 = 1$ ,  $a_2^1 = 0.001$ ,  $a_3^1 = 1$  e  $a_2^0 = 0.001$ , o parâmetro equivalente seria  $a_2^{eq1} = 0.002$ . Se os valores  $a_2^0 = 0.001$  e  $a_2^1 = 0.001$  não corresponderem a uma transformação, mas sim a um erro numérico do método, teremos um acúmulo de erro de  $a_2^{eq} = 0.002$ . Esse processo de acúmulo pode se tornar crítico para transformações com muitas concatenações, pois o acúmulo de erros tem um comportamento semelhante ao de uma progressão aritmética. Este acúmulo de erros pode ser mostrado na Figura 8.



**Figura 8-** Acúmulo de erros nos parâmetros de rotação.

Na seqüência Stefan, nos primeiros 20 quadros existe somente zoom e a translação, nos movimento de fundo. Porém, como mostra a Figura 8, o acúmulo de erros na estimação confunde o método e um movimento de rotação é estimado. Assim foi necessário criar uma forma de evitar esse acúmulo de erros,

pequenos para estimações quadro a quadro, porém significativo com um maior número de quadros. Outra constatação foi de que os termos de rotação são mais sensíveis à concatenação, pois numericamente são menores.

Uma forma de atenuar os efeitos do acúmulo de erros, para os termos de rotação que são mais sensíveis, é a saturação das estimações quadro a quadro,

$$a_{1,2}^n = \begin{cases} 0, & a_{1,2}^n < \text{limiar} \\ a_{1,2}^n, & a_{1,2}^n \geq \text{limiar} \end{cases} \quad (43)$$

Os parâmetros com valores de módulo muito baixo são desconsiderados na estimação e substituídos por zero.

Esse processo atenua um pouco o acúmulo de erro na estimação dos parâmetros, porém não é efetivo para erros que estão presentes em parâmetros com valores acima do limiar de saturação.

Um outro problema da saturação em valores mínimos das transformações é a diminuição da precisão do método. Com este método é necessário que os valores das transformações estejam acima de um certo valor, para serem detectados. Porém não influencia muito nos erros das estimações dos parâmetros, onde são pouco significativos, pois possuem valores pequenos. Além de mostrarem-se pouco perceptíveis, nos quesitos de qualidade do mosaico e dos quadros estimados.

### 3.3. Método de composição do mosaico

Após a obtenção dos parâmetros de movimento, é realizada a etapa de fusão do quadro atual transformado com o mosaico, baseado em partes do trabalho de A. Smolic [8]. A fusão é feita através de uma média ponderada entre o quadro atual transformado e o mosaico construído até o quadro anterior. O cálculo da média ponderada é feito com o auxílio de uma máscara de pesos, calculada para cada quadro  $n$ ,

$$Mosaico_{\text{pesos}}(n) = \begin{bmatrix} \text{peso}_{00} & \dots & \text{peso}_{0n} \\ \vdots & & \vdots \\ \text{peso}_{n0} & \dots & \text{peso}_{nn} \end{bmatrix}. \quad (44)$$

Os pesos da máscara correspondem ao número de *pixels* dos quadros transformados, um de cada quadro compensado, que compõe determinado *pixel* do mosaico,

$$Mosaico_{\text{pesos}}(n, i, j) = \sum_{k=0}^{n-1} ICe(k, i, j). \quad (45)$$

Onde  $ICe(n)$  verifica se um certo *pixel* do mosaico existe no quadro compensado, sendo a função dada por:

$$ICe(n, i, j) = \begin{cases} 1, & I_n(i, j) \in Mosaico \\ 0, & I_n(i, j) \notin Mosaico \end{cases}. \quad (46)$$

O peso do quadro transformado  $I(n)$ , que irá compor o mosaico é unitário, sendo igual a  $ICe(n)$ . A cada novo quadro uma nova máscara de pesos do mosaico é calculada a partir da expressão (45).

O mosaico é calculado, a cada quadro, pela expressão:

$$Mosaico(n, i, j) = \frac{1}{Mosaico_{\text{pesos}}(n, i, j) + ICe(n, i, j)} \left[ \begin{matrix} Mosaico_{\text{pesos}}(n, i, j) \cdot Mosaico(n-1, i, j) \\ + ICe(n, i, j) \cdot I(i, j) \end{matrix} \right] \quad (47)$$

Onde  $Mosaico(n, i, j)$  é o *pixel* do mosaico dos quadros de 0 até  $n$ ,  $Mosaico(n-1, i, j)$  é o *pixel* do mosaico dos quadros de 0 até  $n-1$ .

A média ponderada é utilizada, pois diminui a influência dos objetos que aparecem na frente do fundo em um pequeno número de quadros. Através da

média ponderada, é possível atenuar a influência de *pixels* de objetos com valores bem diferentes na composição do fundo da imagem. Outra razão da utilização da média ponderada é que ela torna possível diminuir a influência da diferença de iluminação na composição do mosaico. Com a média ponderada é possível encontrar o ponto, na média de todos os quadros, que está mais adequado para a minimização dos erros entre o fundo estimado e o fundo real.

A compensação por média ponderada tem como problema não diferenciar o fundo dos objetos, pois não existe uma segmentação prévia. Como não existe a diferenciação entre objeto e o fundo do quadro, objetos com pouco movimento, em relação ao fundo do quadro, acabam compondo excessivamente alguma região do mosaico, borrando o mosaico.

## **3.4. Padrões e critérios de desempenho**

### **3.4.1. Padrão dos vídeos de teste**

Para alguns testes foram utilizadas seqüências de vídeo no formato cif (*Common Intermediate Format*), com freqüência de 30Hz. O padrão possui três componentes, uma luminância e duas crominâncias. A componente de luminância, com tamanho de 352x288 pixels, e as duas crominâncias, com tamanho de 176x144 pixels. As componentes de crominância são decimadas, pois a sensibilidade humana destas componentes é bem menor.

Outros testes também foram feitos com seqüências de vídeo, com freqüência de 30Hz, formadas de várias imagens bitmap de 24 bits. No padrão bitmap são utilizados três componentes R (vermelho), G (verde) e B (azul). Os três componentes com tamanho de 352x288 pixels e com 255 níveis de intensidade.

Para a estimação de movimento foi escolhida são utilizados quadros em níveis de cinza, pela facilidade na conversão. Os quadros em níveis de cinza são formados pela média simples das três componentes R, G e B. Nas seqüências no formato cif é feita, no início do processo, uma conversão dos quadros para o formato RGB. Como o formato cif decima as crominâncias, é necessário, antes da conversão em RGB, interpolar os valores de crominância. Após a conversão para RGB, é feita uma conversão em níveis de cinza, para a estimação de movimento.

### **3.4.2. Ferramentas utilizadas**

Foi desenvolvido um programa para a extração do fundo de uma seqüência de vídeo, utilizando a linguagem de programação C++. A escolha da linguagem C++ foi devido à familiaridade e a facilidade de encontrar bibliotecas, para manipulação dos quadros de vídeo. Outra questão da utilização de C++ é devido ao alto desempenho, comparado com sistemas interpretados como o Matlab® ou Maple®. Para visualização dos resultados foi utilizado o software Matlab®, pois

possui fácil manipulação dos gráficos e não necessita de um processamento muito pesado. Todas as simulações foram feitas em um PC Intel Pentium 4 1.6 GHz com 448 MBytes de memória RAM.

### 3.4.3. Cálculo do PSNR (dB)

Primeiramente o quadro estimado é extraído do mosaico, através de uma transformação geométrica. Após a estimação do quadro, é feito o cálculo do MSE entre o quadro original e o quadro estimado. Através do MSE é calculada a raiz do erro médio quadrático  $e_{rms}$  e por fim a relação sinal de pico ruído (PSNR) [11],

$$\text{PSNR} = 20 \cdot \log \left( \frac{255}{e_{rms}} \right). \quad (48)$$

A relação sinal de pico ruído é uma forma de normalizar o erro médio quadrático entre o quadro original e o quadro estimado. Para o pico do sinal é utilizado o valor de 255, que neste caso é o máximo valor que o sinal pode atingir e o ruído é representado pelo erro médio quadrático.

### 3.4.4. Quadro estimado

O quadro é estimado a partir do mosaico e dos parâmetros de movimento global [12]. Como cada quadro possui seus respectivos parâmetros de movimento em relação à referência, primeiro quadro para este trabalho, é possível transformar o mosaico em um quadro. Esta operação se comporta como o inverso da composição do mosaico.

O quadro reconstituído tem os *pixels* com influência de todos os quadros, que compõe o trecho mapeado, para a reconstrução. A influência de vários quadros, para compor o mosaico pode gerar, em alguns casos, erros grandes na estimação de um quadro específico.

### **3.5. Considerações finais do capítulo**

Neste capítulo foi proposto um método de estimação do mosaico, baseado na filtragem dos vetores de movimento. O método utiliza critérios de filtragem dos vetores de movimento baseados na homogeneidade e na extração dos vetores de objetos.

O método proposto neste capítulo será empregado no próximo capítulo para testes de desempenho em relação à qualidade. Também serão apresentados no próximo capítulo os mosaicos, utilizando a técnica proposta neste capítulo.

## 4. Resultados

Neste capítulo serão apresentados resultados de desempenho do método, na estimação do mosaico, e comparativos com outros métodos. Outros fatores que serão observados nos resultados serão o desempenho na estimação dos parâmetros de movimento global e o papel da interpolação na qualidade do mosaico.

Para realizar a análise da qualidade do mosaico estimado foram feitos testes, através da reconstrução dos quadros a partir do mosaico. Com os quadros reconstruídos são calculados os PSNR.

O estimador de parâmetros de movimento global foi testado para situações de translação pura, porém com alguns vetores de movimento espúrios em quantidades razoáveis, que podem gerar erros de estimação. Neste teste pode ser verificado o efeito da saturação, nos valores de rotação, em uma seqüência de vídeo.

Na análise da estimação dos parâmetros de movimento global é feito um comparativo entre os parâmetros estimados e os parâmetros reais. Os parâmetros reais são conhecidos somente das seqüências de vídeo geradas artificialmente, não sendo conhecidos em vídeos naturais. Esse teste é feito com a intenção de conhecer os limites do método na estimação de movimento global.

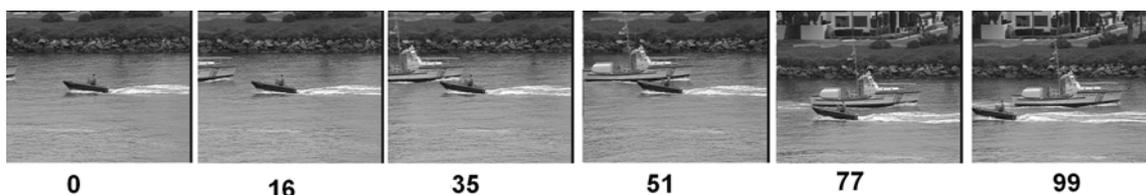
Foram realizados testes para verificar os limites da qualidade da reconstrução dos quadros. Estes testes servem também para analisar os efeitos da interpolação na qualidade do quadro reconstruído. Pois foi observada uma degradação na qualidade dos quadros reconstruídos com os parâmetros de movimento ideais, não perceptíveis visualmente, mas que influenciarão no resultado final.

Para os testes foram utilizadas onze seqüências de vídeo, onde três são vídeos naturais e nove de vídeos artificiais. Das seqüências de vídeo natural foram utilizadas as seqüências Stefan.cif, bus.cif e coast\_guard.cif, seqüências clássicas para teste de vídeo. Já nos vídeos artificiais foram utilizadas sete seqüências

Lenna\_zoom, Luana\_zoom, Peixe\_zoom, seqüências somente com zoom. Lenna\_rotação, Luana\_rotação, Peixe\_rotação, seqüências somente com rotação. Peixe\_translação, seqüência somente com translação, e Lenna\_completa, com rotação, zoom e translação.

## 4.1. Mosaicos

Nesta seção serão mostrados os mosaicos, resultado final do processo, estimados com o método desenvolvido. O primeiro mosaico é referente à seqüência coast\_guard, usando 100 quadros. Alguns quadros da seqüência coast\_guard são mostrados na Figura 9.



**Figura 9**-Alguns quadros da seqüência coast\_guard.

Na Figura 10 é apresentado um mosaico gerado pelo método proposto. É perceptível o borramento no centro do mosaico, referente aos dois barcos que estão presentes na seqüência. Também são visíveis algumas manchas, parecidas com linhas pretas. Estas linhas são da própria seqüência e estão presentes na lateral direita do vídeo. Outra razão do borramento é o movimento da água durante toda a seqüência de vídeo. Os outros erros podem ser atribuídos ao próprio método de estimação.

Porém, visualmente, a qualidade do mosaico é bastante aceitável, sendo bastante parecida com a seqüência de vídeo. É possível distinguir até detalhes pequenos, como as duas pessoas conversando em baixo do toldo e o pequeno arbusto em frente à porta da casa.

No quesito compensação, o comportamento foi aceitável também. Os dois objetos que estavam à frente do fundo, os dois barcos, sumiram, não sendo possível distinguir quais objetos estavam à frente do fundo.



**Figura 10-** Mosaico de 100 quadros da seqüência coast guard.

A Figura 11 mostra alguns quadros da seqüência Stefan, que são utilizados para construir o mosaico como é mostrado na Figura 12. Para construir este mosaico foram necessários 100 quadros da seqüência Stefan.



**Figura 11-** Alguns quadros da seqüência Stefan.

Em termos de qualidade este mosaico é menos nítido do que o mosaico da seqüência coast\_guard. É visível no centro do mosaico, um borramento gerado pelo objeto, o jogador, que estava à frente do fundo. Também é visível um borramento na arquibancada ocasionado, entre outros fatores, pela movimentação

das pessoas na arquibancada. São perceptíveis também no canto direito linhas escuras, que estão presentes na seqüência e que geram borramento no mosaico.

A compensação não teve tanto êxito nesta seqüência quanto na anterior. Pois, o objeto à frente do fundo ficava durante vários quadros em uma única posição. A bola no canto esquerdo também não foi compensada, pois o número de quadros em que a bola aparece naquela posição é significativo, em relação ao número de quadros que compõe aquela região do mosaico.

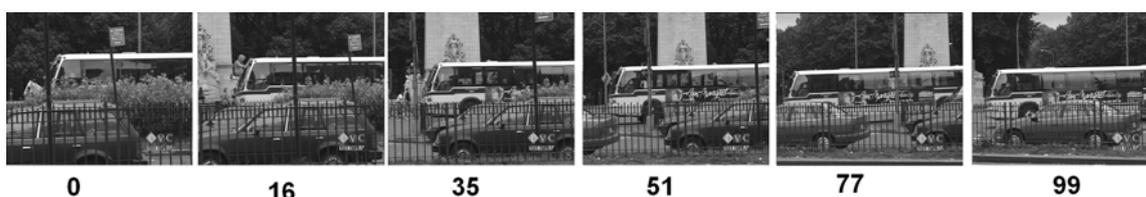
Podem-se citar alguns pontos positivos do mosaico da seqüência Stefan. A linha do muro atrás da quadra está praticamente reta, com poucas distorções. Existe pouca distorção de tamanho no decorrer da construção do mosaico, pois os dois homens que estão encostados no muro possuem tamanhos semelhantes. Várias partes dos letreiros estão visíveis e com proporções de tamanhos adequados em relação ao muro. As linhas da quadra estão praticamente retas, somente com pequenas distorções. Várias pessoas estão bastante visíveis na arquibancada, principalmente do lado direito e do lado esquerdo do mosaico.

De uma maneira geral a qualidade do mosaico da seqüência Stefan ficou razoável. Porém, alguns erros de estimação ocorreram em função da natureza do método de estimação.



**Figura 12-** Mosaico de 100 quadros da seqüência Stefan.

O mosaico da seqüência bus é mostrado na Figura 14. Comparando com os dois casos anteriores, esta é a pior situação, pois é possível ver neste mosaico, que alguns problemas de estimação de movimento global e de compensação ocorreram. A estimação de movimento global, neste caso, foi bastante prejudicada, pois o objeto é grande, se comparado com o tamanho do quadro, como pode ser visto na Figura 13 que mostra alguns quadros da seqüência bus.



**Figura 13-** Alguns quadros da seqüência bus.

Com objetos grandes à frente do fundo ocorre, no método proposto, uma confusão em distinguir o que é objeto e o que é fundo, gerando problemas para encontrar o movimento global. Esta confusão ocorre devido ao critério de distinção entre vetores de movimento de objetos e de fundo, que neste trabalho é dependente da quantidade de vetores de movimento. Como a quantidade dos vetores de movimento do ônibus é significativa, em relação a quantidade total dos vetores de movimento do quadro, o movimento do ônibus é confundido com o movimento global. Um dos efeitos da má estimação do movimento global é perceptível na estimação da translação. Pois o monumento no centro e o carro da direita aparecem em duplicidade no mosaico. O outro problema que ocorre devido a objetos grandes, a compensação, também é crítico nesta seqüência. Neste caso o objeto é grande, além de aparecer em muitos quadros, o que torna impossível, para o método proposto, compensá-lo.



**Figura 14-** Mosaico de 100 quadros da seqüência Bus.

O mosaico da seqüência Lenna\_zoom é apresentado na Figura 16. Para construir este mosaico foram utilizados 42 quadros da seqüência Lenna\_zoom. Alguns quadros da seqüência Lenna\_zoom são mostrados na Figura 15.



**Figura 15-** Alguns quadros da seqüência Lenna\_zoom.

Como será mostrada pelo desempenho da reconstrução dos quadros e da estimação dos parâmetros de movimento global, a qualidade do mosaico é muito boa. O mosaico apresenta um borramento bem leve e as dimensões estão compatíveis com o quadro inicial, que é o quadro de referência para as transformações. Sendo que um dos fatores que auxiliaram no bom desempenho é a quantidade reduzida de quadros, que é a metade dos quadros da seqüência bus.

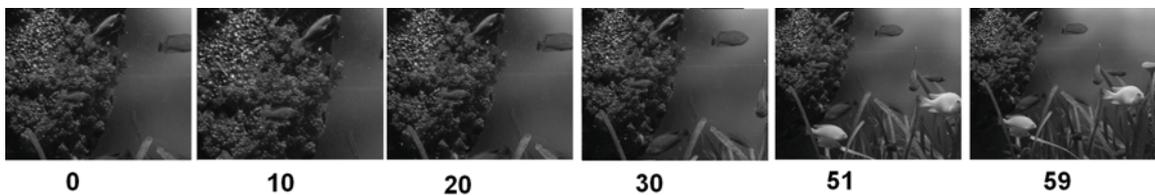


**Figura 16-** Mosaico da seqüência Lennazoom.

O mosaico da seqüência Luana\_zoom, Figura 19, e Peixe\_zoom, Figura 20, também estão com uma qualidade satisfatória. Como na seqüência anterior o grau de borramento nas duas seqüências é pequeno. Alguns trechos das seqüências Luana\_zoom e peixe\_zoom são mostrados nas Figura 17 e Figura 18 respectivamente.



**Figura 17-** Alguns quadros da seqüência Luana\_zoom.



**Figura 18-** Alguns quadros da seqüência peixe\_zoom.



**Figura 19-** Mosaico da seqüência Luana\_zoom.



**Figura 20-** Mosaico da seqüência peixe\_zoom.

A qualidade visual dos mosaicos, das seqüências de zoom, acompanha os valores dos PSNR estimados. E também comprova a eficácia do método para estimação de movimento global, para seqüências de zoom.

As seqüências de rotação são as mais críticas para o método. Como comprovam os resultados da reconstrução dos quadros a partir do mosaico. Os mosaicos resultantes estão extremamente borrados, pois os parâmetros de movimento global estão mal estimados.

A seqüência Lenna\_ rotação como pode ser visto na Figura 21, que mostra alguns quadros da seqüência, gira em torno do ponto de origem lentamente. Pois demora 147 quadros para atingir uma rotação de aproximadamente 180°.



**Figura 21-**Alguns quadros da seqüência Lenna\_ rotação.

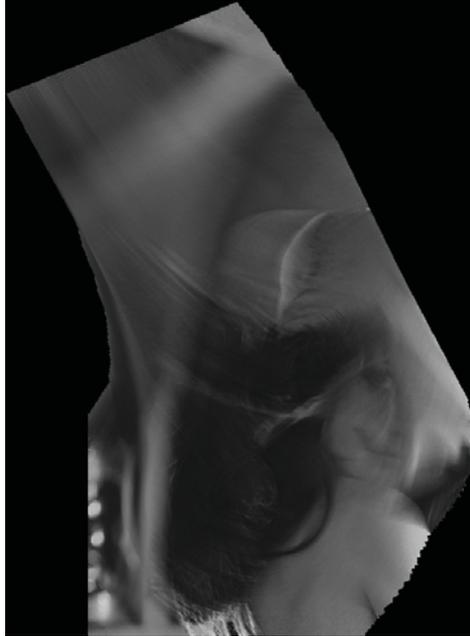
Através da trajetória do ponto central da seqüência Lenna\_ rotação, que é mostrada na Figura 22, é possível ver que a trajetória é suave, variando constantemente.



**Figura 22-** Trajetória do ponto central do primeiro quadro da seqüência Lenna\_ rotação.

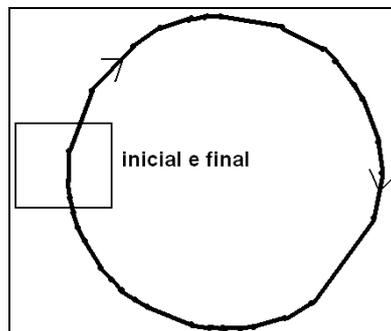
No mosaico da seqüência Lenna\_ rotação, Figura 23, é possível distinguir certas formas, na região inferior do mosaico. Mas na região superior do mosaico o

borramento é excessivo, não sendo possível à distinção de algumas partes, que estão presentes na seqüência Lenna\_ rotação, como o chapéu e o nariz da Lenna.



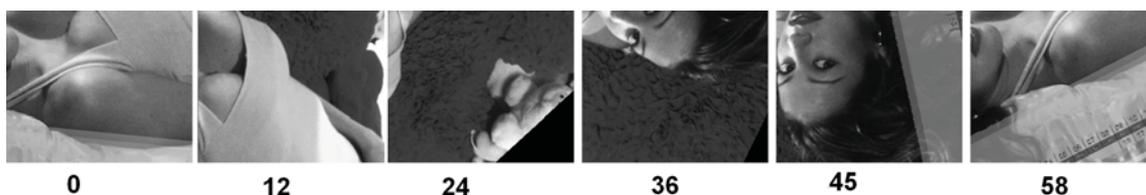
**Figura 23-** Mosaico da seqüência Lenna\_ rotação.

Os três mosaicos das seqüências de rotação apresentam problemas graves, porém no mosaico Luana\_ rotação, Figura 26, eles são mais críticos. Neste caso, como a estimação de movimento global é muito ruim, ocorreu um borramento muito grande. Um dos motivos são as variações muito rápidas de direção,  $360^\circ$  em 58 quadros, que podem ser observadas na trajetória do pixel central do quadro inicial, Figura 24.

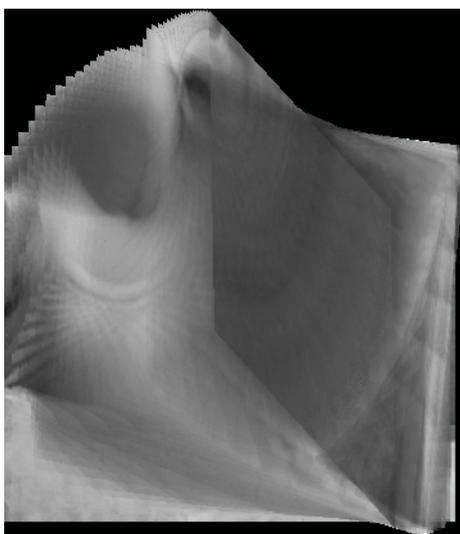


**Figura 24-** Trajetória do ponto central do primeiro quadro da seqüência Luana\_ rotação.

Desta forma, neste mosaico não é possível fazer a distinção de algumas formas, como na situação anterior, como pode ser visto na comparação entre a Figura 25, que mostra alguns trechos da seqüência Luana\_ rotação, e o mosaico da Figura 26.



**Figura 25-**Alguns quadros da seqüência Luana\_ rotação.



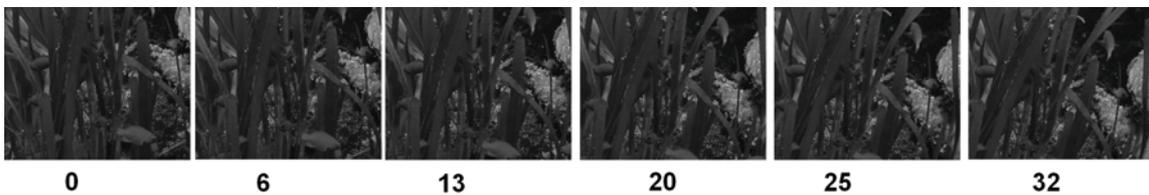
**Figura 26-** Mosaico da seqüência Luana\_ rotação.

O mosaico da seqüência peixe\_ rotação, Figura 27, visualmente obteve um resultado melhor do que as outras duas seqüências de rotação.



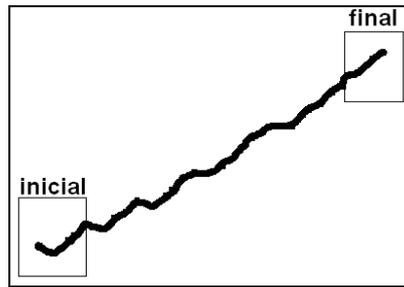
**Figura 27-** Mosaico da seqüência peixe\_rotação.

O borramento é grande ainda, se comparado com a Figura 28 que apresenta alguns trechos da seqüência peixe\_rotação, porém é possível distinguir os peixes e as plantas, em praticamente todas as regiões.



**Figura 28-** Alguns quadros da seqüência peixe\_rotação.

Um dos possíveis motivos do relativo sucesso da seqüência peixe\_rotação é devido a sua variação mais lenta do que as duas outras seqüências. Esta variação mais lenta pode ser observada pela Figura 29.



**Figura 29-** Trajetória do ponto central do primeiro quadro da seqüência peixe\_rotação.

O mosaico da seqüência Lenna\_completa, Figura 30, apresenta um resultado intermediário, em termos de qualidade, entre as seqüências de zoom e as seqüências de rotação.



**Figura 30-** Mosaico da seqüência Lenna\_completa.

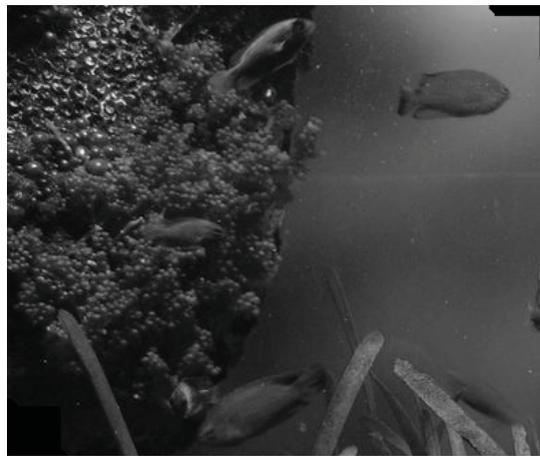
A Figura 31 mostra alguns quadros da seqüência Lenna\_completa.



**Figura 31-** Alguns quadros da seqüência Lenna\_completa.

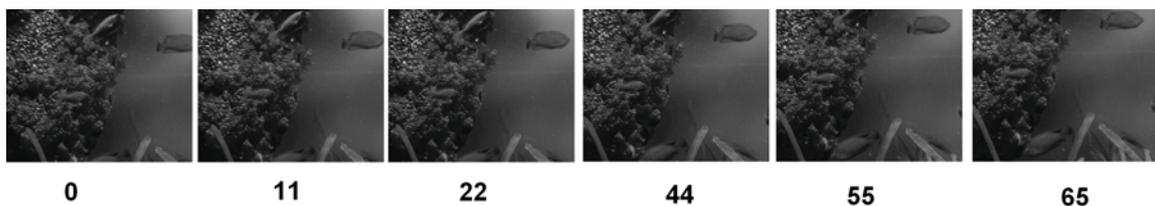
Neste caso o borramento é bem mais leve do que o das seqüências de rotação, porém é maior do que nas seqüências de zoom. Visualmente é possível distinguir as formas, além de regiões com uma qualidade satisfatória, como o canto esquerdo do mosaico, em que o borramento é bem pequeno. Porém em alguns pontos existe duplicidade de partes, como os olhos e o ombro direito da Lenna que aparecem em duplicidade.

O mosaico da seqüência peixe\_translação, Figura 32, é o que possui maior qualidade visual.



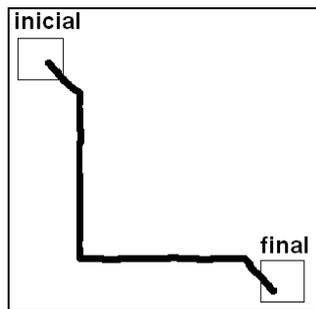
**Figura 32-** Mosaico da seqüência peixe\_translação.

Alguns quadros da seqüência peixe\_translação são mostrados na Figura 33.



**Figura 33-** Alguns quadros da seqüência peixe\_translação.

Mesmo com variações de direção abruptas, que podem ser vistas pela trajetória do ponto central do quadro inicial na Figura 34, o comportamento do estimador de movimento foi satisfatório.



**Figura 34**-Trajetória do ponto central do primeiro quadro da seqüência peixe\_translação.

Não é perceptível nenhum borrramento em todas as regiões do mosaico. Isto ocorre devido à excelente estimação do movimento global, no caso da translação, e da composição ser adequada para esta situação, seqüência sem objetos na frente.

## **4.2. Relação entre a qualidade da compensação, zonas de atividade de movimento e padrões de compensação**

A intenção desses testes é verificar em que pontos da seqüência de vídeo as estimações são mais difíceis e quais as razões destas dificuldades. Será tentado relacionar estas dificuldades nas estimações dos quadros, a partir do mosaico, com os trechos com maior atividade e com padrões de compensação. Para encontrar os quadros com maior atividade de movimento e para obter um dos padrões de qualidade, foi feita a diferença entre os quadros consecutivos. O outro padrão de qualidade de reconstrução deve levar em conta a dificuldade de reconstruir os quadros, sendo utilizada para isso uma compensação quadro a quadro. Na compensação quadro a quadro é feita inicialmente uma estimacão de movimento global entre os quadros consecutivos. Para a estimacão do movimento global é utilizado um método semelhante ao que utilizamos para criar o mosaico, porém não levando em conta o movimento de quadros passados. Após a estimacão dos parâmetros de movimento global, é feita a transformacão do quadro posterior, com os parâmetros de movimento global estimados, seguido da subtração entre o quadro posterior transformado e o quadro anterior.

Serão analisados nestes gráficos, visualmente, principalmente dois quesitos. A sensibilidade da estimacão em função da atividade da seqüência e também a qualidade em relação aos padrões de compensação. A diferença quadro a quadro, a compensação pela reconstrução a partir do mosaico e a compensação quadro a quadro serão, em alguns pontos na próxima seção, citados como Q a Q, Mosaico e Q a Q compensada respectivamente.

#### 4.2.1. Análise dos resultados

Na comparação entre os PSNR, Figura 35, é percebido que a compensação com mosaico tem um desempenho inferior às outras duas. Porém no início da seqüência bus, onde a translação é mais visível, o desempenho da compensação por mosaico é melhor que a pura diferença entre os quadros. Nos locais de maior atividade, onde a diferença quadro a quadro é maior e o PSNR é menor, a compensação através do mosaico possui valores de máximos locais.

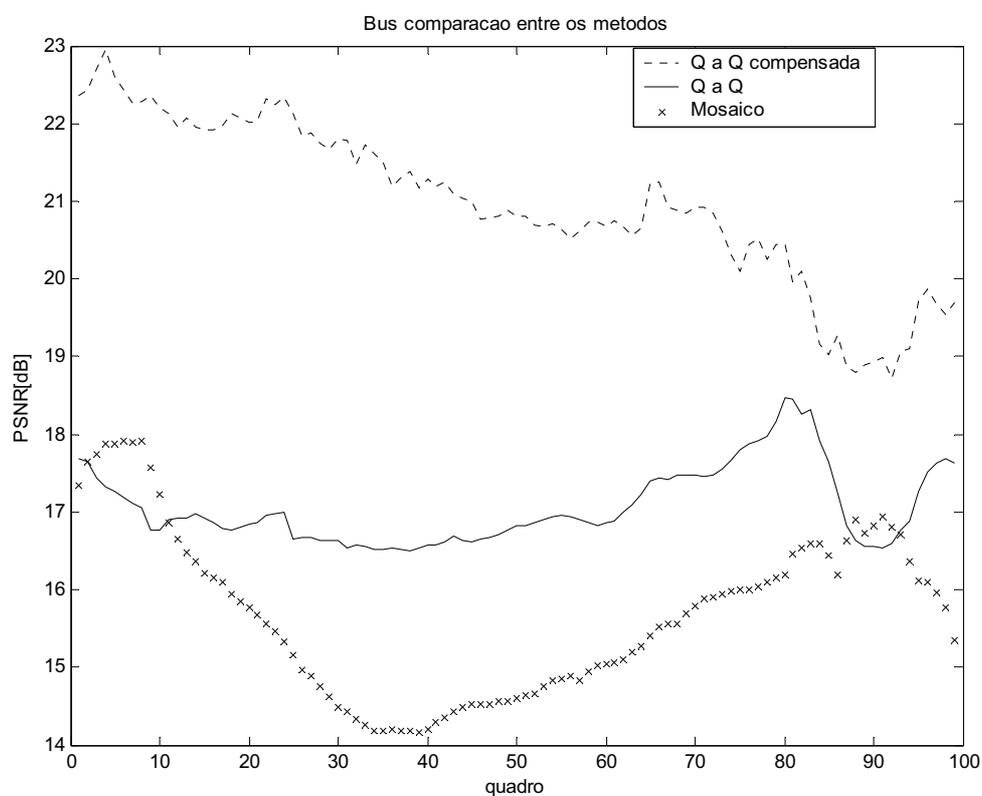
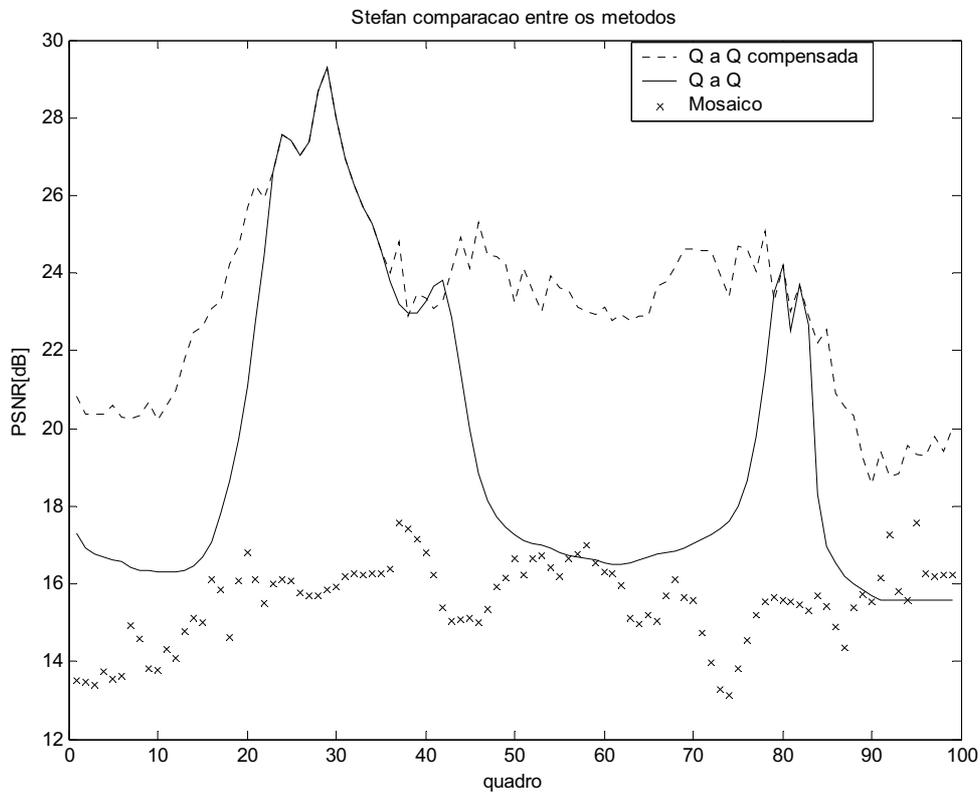


Figura 35- Seqüência bus, comparação entre os métodos.

Comparando os PSNR da Figura 36, referentes à seqüência Stefan, é percebido que o desempenho da compensação pelo mosaico não teve um comportamento muito satisfatório. Porém, como foi visto no gráfico da seqüência bus existe um comportamento melhor nas regiões do gráfico com alta atividade de movimento. Outro ponto verificado, é que o comportamento da compensação pelo mosaico é menos sensível a grandes variações entre um quadro e outro do que uma compensação quadro a quadro. Como o gráfico mostra, a variação do PSNR na compensação quadro a quadro é de aproximadamente 10dB enquanto que a variação da compensação por mosaico é de aproximadamente 4dB.



**Figura 36-** Seqüência Stefan, comparação entre os métodos.

Na Figura 37, referente à seqüência coast guard, é verificado novamente um comportamento da compensação quadro a quadro, melhor do que uma compensação por mosaico. Porém para regiões de grande atividade, a compensação por mosaico é mais eficaz, somente no quadro setenta. Nesta seqüência a compensação por mosaico é extremamente mais suave que as outras duas compensações, com uma variação máxima de 2 dB.

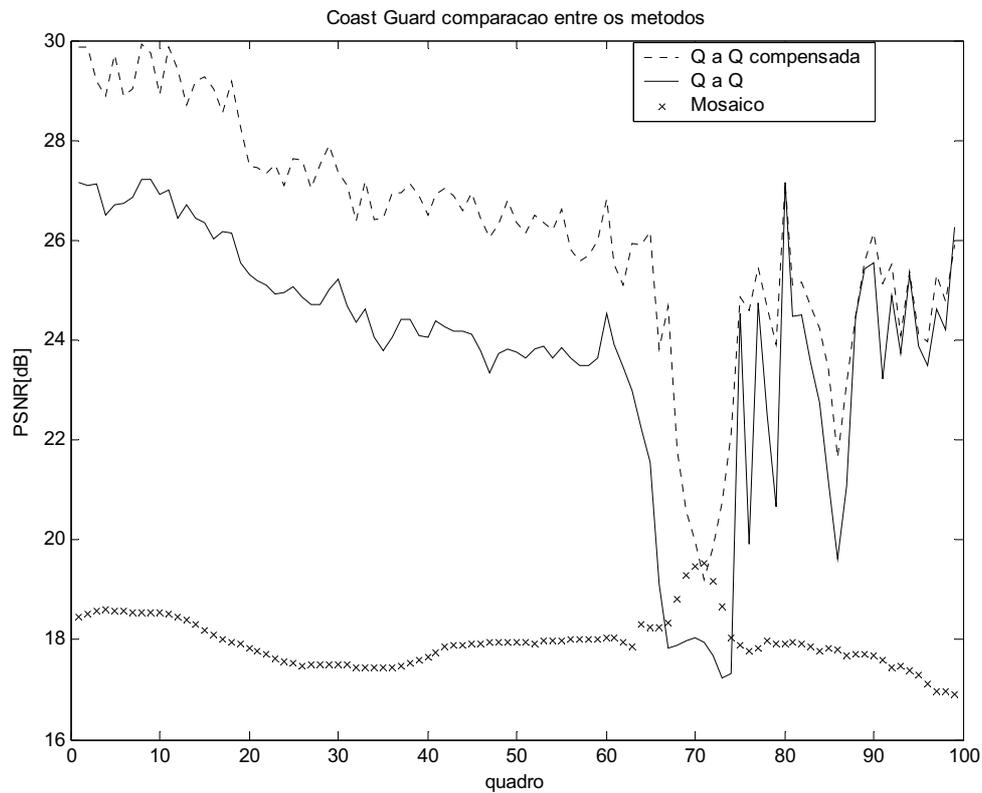
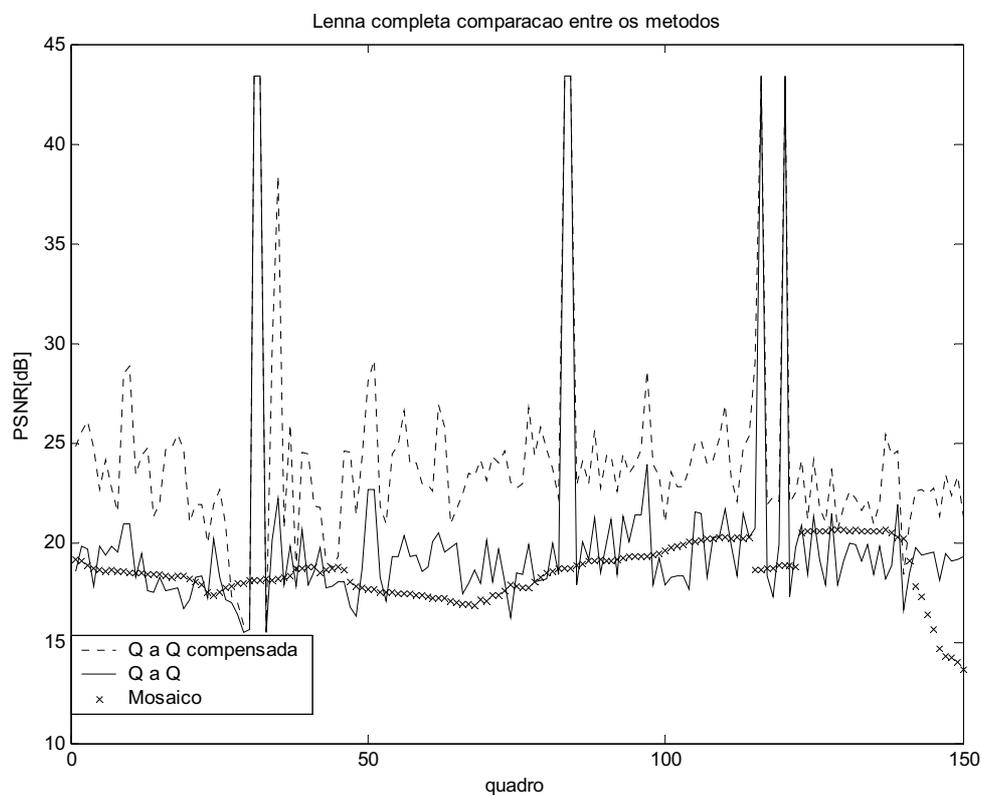


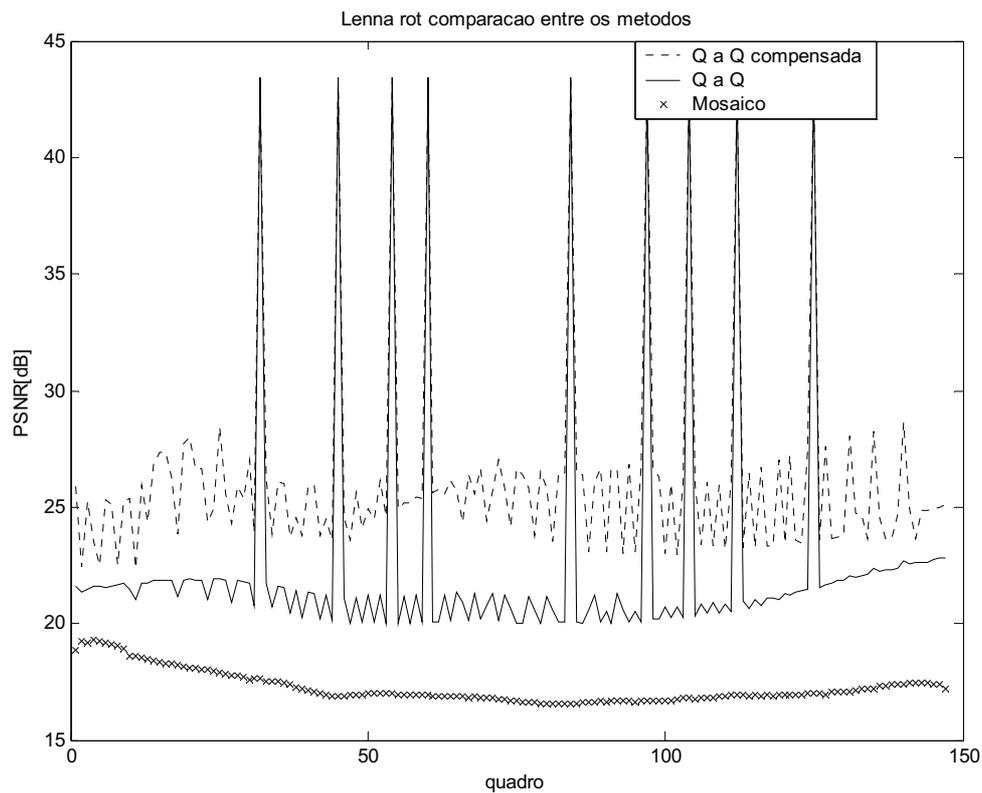
Figura 37- Seqüência coast\_guard, comparação entre os métodos.

Novamente o PSNR do mosaico, Figura 38, tem um comportamento mais suave em relação à estimação quadro a quadro e a simples diferença entre os quadros. Outra questão, é que a compensação por mosaico, neste caso, tem um comportamento suave mesmo em zonas de grande atividade. Entre o quadro 27 e 30 a compensação por mosaico é mais eficiente. Os picos correspondentes a mais de 40dB de PSNR são de PSNR infinito, pois essas são as partes da seqüência em que não existe movimento. Porém, nesta situação ainda não temos um comportamento ideal para compressão de vídeo.



**Figura 38-** Seqüência Lenna\_completa, comparação entre os métodos.

O gráfico da Figura 39 retrata uma situação limite para a compensação através do mosaico, pois é um caso de grande rotação. Neste exemplo a qualidade final da estimação ficou bem ruim. Outro problema neste caso é a questão da quantidade de quadros. Como a estimação não é muito boa existe um excessivo acúmulo de erros na estimação, tornando a qualidade em relação às outras compensações ruins. Pois as outras compensações não têm sua qualidade diminuída em função da quantidade de quadros. Outra questão que pode ser vista nesta seqüência, da mesma forma que nos outros gráficos, é que o comportamento é bastante suave em relação à diferença quadro a quadro e à compensação quadro a quadro.



**Figura 39-** Seqüência Lenna\_ rotação, comparação entre os métodos.

Nestes outros dois casos de rotação, retratados na Figura 40 e na Figura 41, é possível ver o efeito do acúmulo de erros. Nestas duas seqüências as quantidades de quadros são bem menores do a seqüência anterior. Sendo os parâmetros de zoom bastante sensíveis a quantidade de quadros da seqüência. É possível observar uma melhora considerável na qualidade relativa do PSNR, nas duas próximas seqüências. Outra questão observável nestes gráficos é o comportamento da compensação feita com o mosaico em situações de mínimos locais. Onde a estimação tem um comportamento menos hostil em relação às outras compensações. Porém, mesmo com a melhora, em relação ao caso anterior, a rotação é uma situação limite para a qualidade do mosaico.

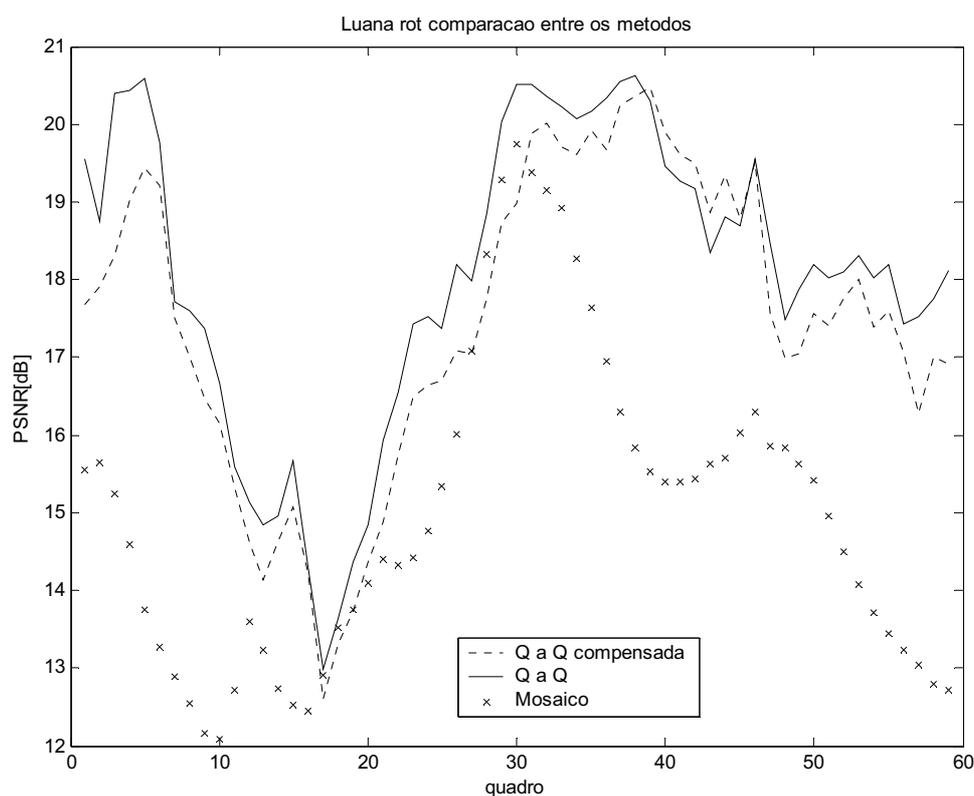
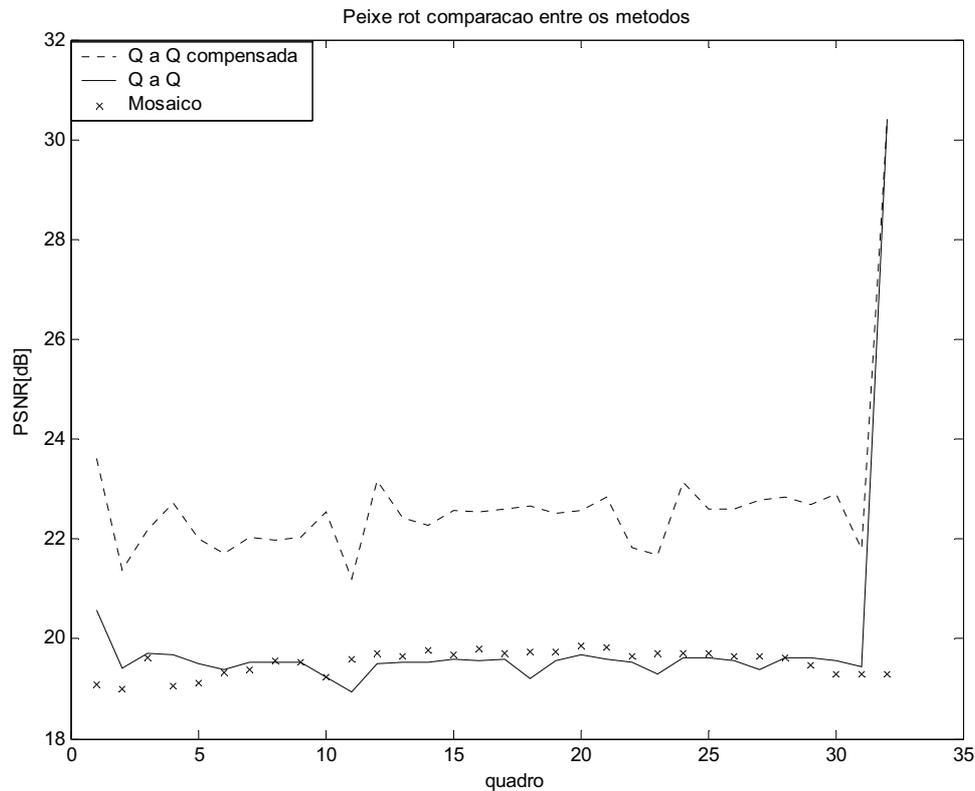
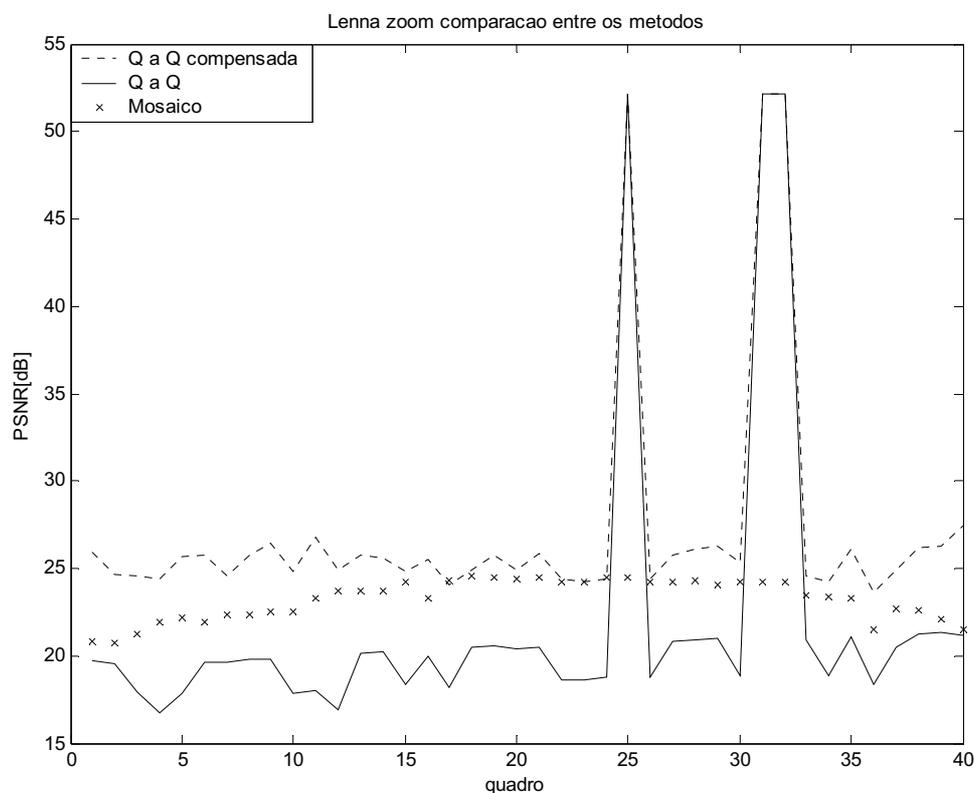


Figura 40- Seqüência Luana\_ rotação, comparação entre os métodos.



**Figura 41-** Seqüência peixe\_rotaçã, comparação entre os métodos.

A próxima seqüência, Figura 42, é uma seqüência de zoom, onde a compensação por mosaico apresenta um comportamento eficaz. Nesta seqüência a compensação por mosaico possui valores relativos de PSNR maiores que a diferença quadro a quadro. E em alguns casos a compensação por mosaico tem um PSNR com valores semelhantes aos PSNR da compensação quadro a quadro. A compensação por mosaico também possui um comportamento suave, modificando pouco os seus valores, quando a atividade de movimento aumenta muito. Os dois picos que esta seqüência apresenta são relativos aos quadros sem movimentação, com PSNR infinito, que foram saturados em 52dB.



**Figura 42-** Seqüência Lenna\_zoom, comparação entre os métodos.

As duas outras seqüências também são seqüências de zoom, Luana\_zoom, Figura 43, e peixe\_zoom, Figura 44. Nestas seqüências a compensação por mosaico, possui um desempenho satisfatório em termos de qualidade em relação à diferença entre os quadros. E em alguns casos, na seqüência peixe\_zoom no quadro três, desempenho melhor. Como nos outros gráficos, novamente pode ser observado que o comportamento do mosaico é pouco modificado pela atividade de movimento, mais visível no caso da seqüência Luana\_zoom. Porém, é necessária uma certa quantidade de quadros para realmente tornar o PSNR suave. Como pode ser visto, na seqüência peixe\_zoom que mesmo tendo um comportamento satisfatório, o PSNR é bastante influenciado pela atividade de movimento. Os dois picos do gráfico do Luana\_zoom são pontos sem movimento, saturados em 66 dB para melhor visualização do gráfico.

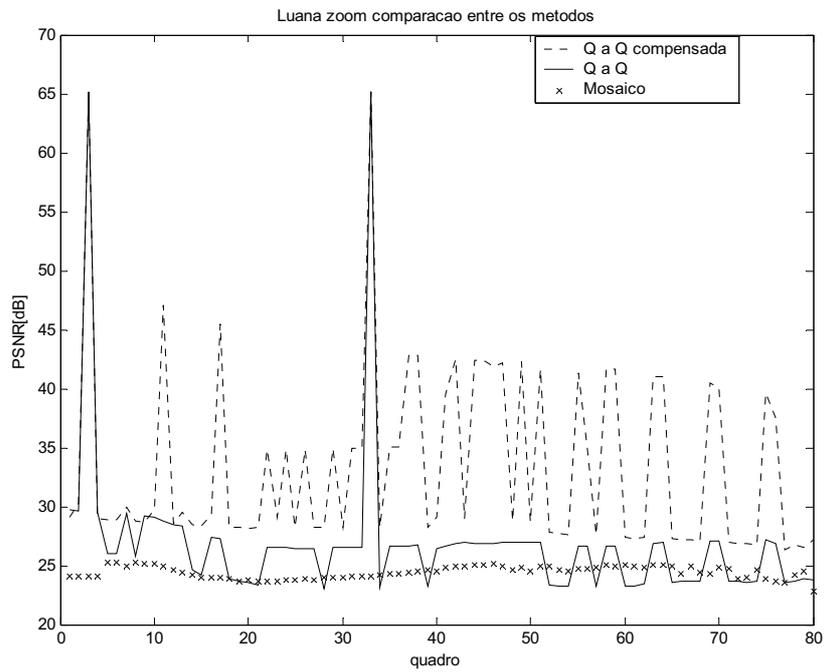


Figura 43- Sequência Luana\_zoom, comparação entre os métodos.

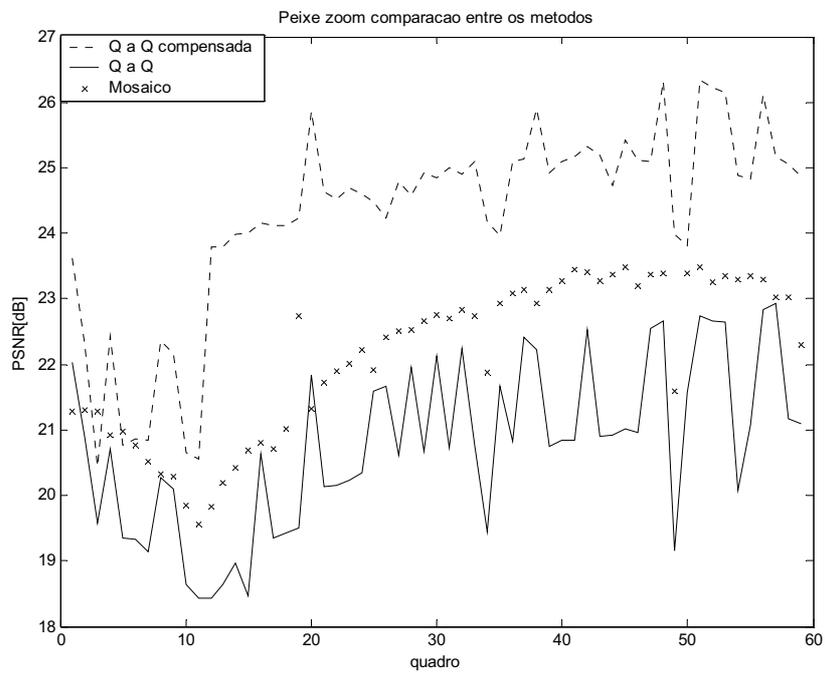
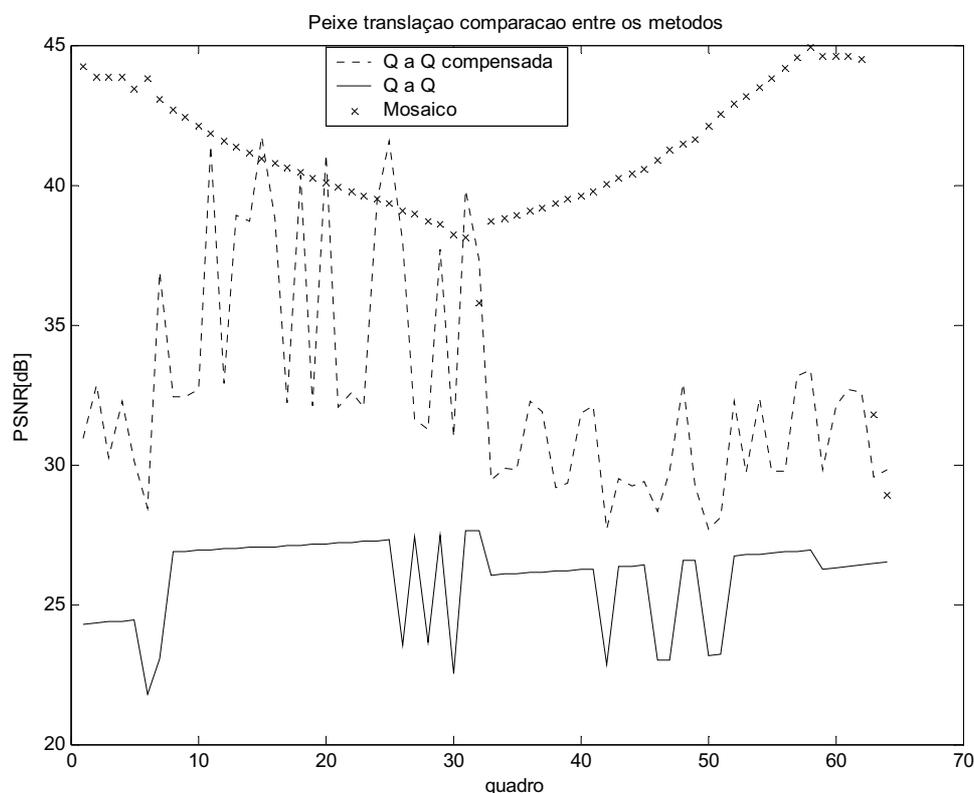


Figura 44- Sequência peixe\_zoom, comparação entre os métodos.

Este é um caso, Figura 45, em que uma compensação por mosaico é mais adequada, a translação. Nesta seqüência, a peixe\_translação, a compensação por mosaico é mais eficaz que uma compensação quadro a quadro. Outra questão importante é a suavidade novamente, nesta seqüência é visível que a compensação por mosaico é pouco influenciada pela atividade da seqüência. Portanto, tem um comportamento muito mais suave.



**Figura 45-** Seqüência peixe\_translação, comparação entre os métodos.

Percebe-se através deste testes que a compensação por mosaico é menos eficaz, na maioria das vezes, que uma compensação quadro a quadro. Porém, ela é extremamente mais estável se comparada com uma compensação quadro a quadro. E atua bem em situações que a compensação quadro a quadro tem seu pior desempenho. Também é visível que este método para compensação por

mosaico é pouco eficaz em situações de rotação, principalmente. Mas a compensação por mosaico mostrou-se bastante eficaz, em termos de qualidade e sensibilidade, para translações. Para zoom o método teve um comportamento intermediário, não foi tão bom quanto em situações de translação, mas não foi tão ruim quanto na rotação. Nas situações de zoom a sensibilidade foi pequena, em relação à atividade, porém em termos de qualidade os resultados não foram tão bons quanto na translação. Para os vídeos naturais foi verificado que em algumas situações de mínimos, das outras duas compensações, houve picos da compensação por mosaico. Ratificando que a compensação por mosaico atua bem em situações que a compensação quadro a quadro não atua bem.

### **4.3. Comparação entre os vídeos naturais**

Os gráficos de PSNR dos vídeos naturais da Figura 46 retratam a reconstrução dos quadros, a partir do mosaico construído, utilizando o método proposto. Neste comparativo entre as três seqüências de vídeo natural, é visível que a seqüência `coast_guard` tem um desempenho bem melhor. Alguns fatores são importantes para o bom desempenho da seqüência `coast_guard`. Primeiro, a seqüência possui, como movimento principal, na maioria dos quadros, translação. Segundo, os objetos que estão na frente do fundo são pequenos, borrando pouco o mosaico no momento da composição. Já a seqüência `bus` possui um objeto grande, o ônibus, em quase todos os quadros prejudicando desde a composição até a estimação dos parâmetros de movimento. Em contrapartida, a seqüência `Stefan`, que possui movimentos até mais complexos que a seqüência `bus`, tem um desempenho médio melhor. Isso se deve à facilidade da composição, pois os objetos à frente do fundo são pequenos. Como objetos da seqüência `Stefan` são pequenos, também é facilitada a estimação dos parâmetros, pois influenciam pouco no comportamento do movimento global.

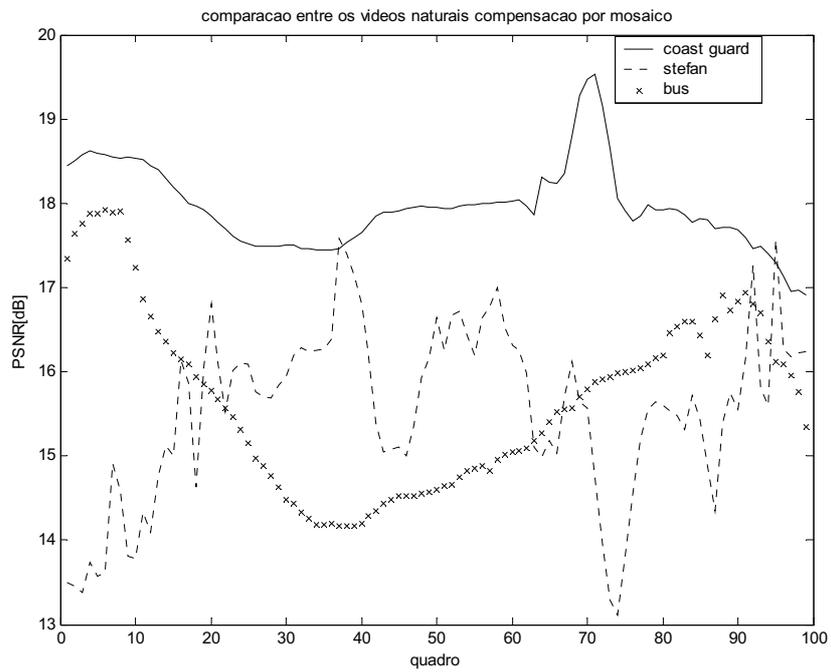


Figura 46- Comparação entre o PSNR dos vídeos naturais, compensação por mosaico .

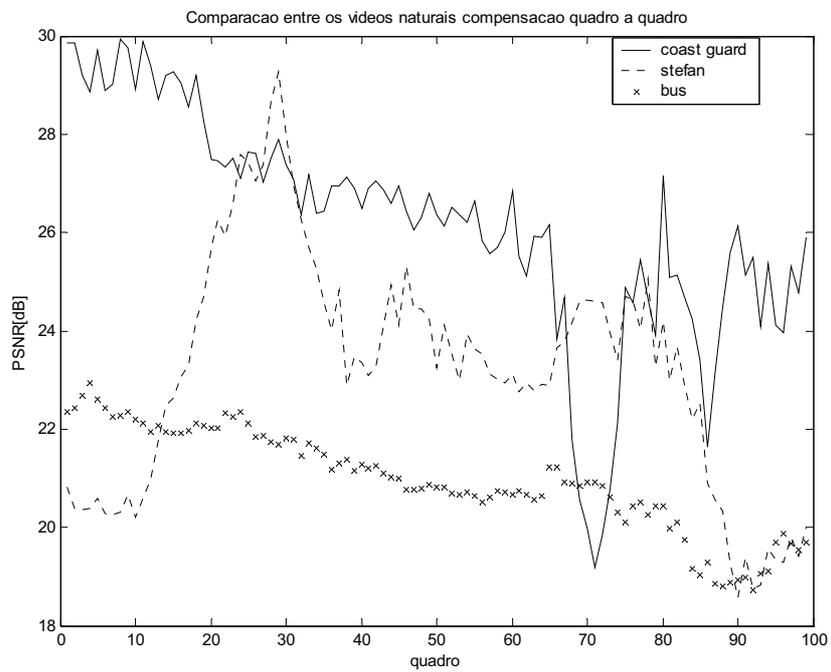
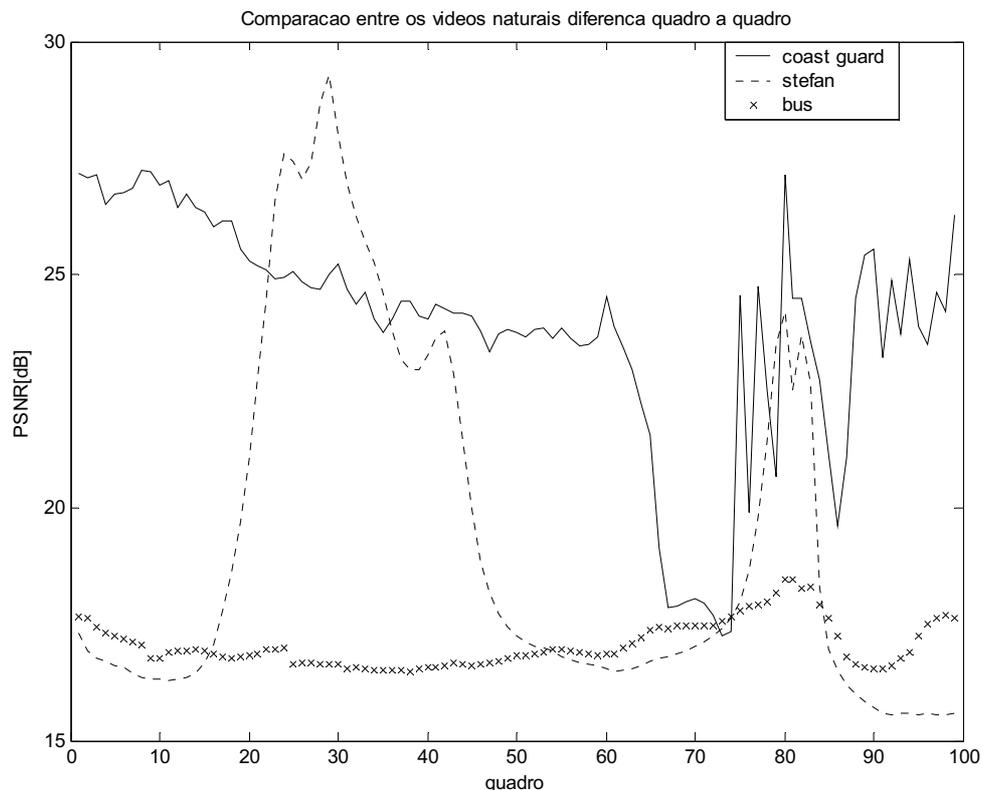


Figura 47- Comparação entre o PSNR dos vídeos naturais, compensação quadro a quadro .

Na compensação quadro a quadro, Figura 47, é observável que a seqüência `coast_guard` possui um desempenho melhor na qualidade de reconstrução. Porém em alguns quadros a seqüência `coast_guard` possui variações abruptas de qualidade, que não ocorrem na seqüência `bus`, por exemplo. Também é observável, em alguns quadros, que a seqüência `Stefan` possui uma qualidade de reconstrução melhor do que a seqüência `coast_guard`. Este comportamento da qualidade de reconstrução por compensação quadro a quadro está diretamente relacionado com a variação entre os quadros subseqüentes.

Através da Figura 48 é possível ver que a variação entre um quadro e seu subseqüente é maior na seqüência `bus`. Esta maior variação tem reflexo direto na qualidade do mosaico, pois a seqüência `bus`, entre os vídeos naturais, obteve a pior qualidade de reconstrução. Já a seqüência `coast_guard`, que possui uma variação média menor entre um quadro e seu subseqüente, possui melhor qualidade de reconstrução entre as três seqüências. Dessa forma é possível ver que a qualidade relativa da reconstrução dos quadros a partir do mosaico está diretamente relacionada com a variação entre os quadros subseqüentes. Isto decorre da utilização da estimação de movimento local entre quadros, através de procura por blocos, para encontrar o movimento global do quadro em relação ao mosaico.



**Figura 48-** Comparação entre o PSNR dos vídeos naturais, sem compensação.

#### 4.4. Influência da interpolação na qualidade do mosaico

A princípio não era imaginado que a interpolação teria efeitos muito visíveis, na qualidade do mosaico. Porém foi verificado, através de testes com os vídeos artificiais, que a interpolação tem um efeito na qualidade do mosaico e da reconstrução dos quadros a partir do mosaico. Portanto, mesmo com uma estimação perfeita dos parâmetros de movimento teremos erros na reconstrução dos quadros. Os erros de reconstrução ocorrem pelo fato dos *pixels* do mosaico, por muitas vezes, não representarem somente um *pixel* do quadro. E no momento da reconstrução, um *pixel* do quadro reconstruído, também por muitas vezes, não é composto de um único *pixel* do mosaico. Esse fato da não correspondência direta entre o *pixel* do quadro e o *pixel* do mosaico gera erros de reconstrução.

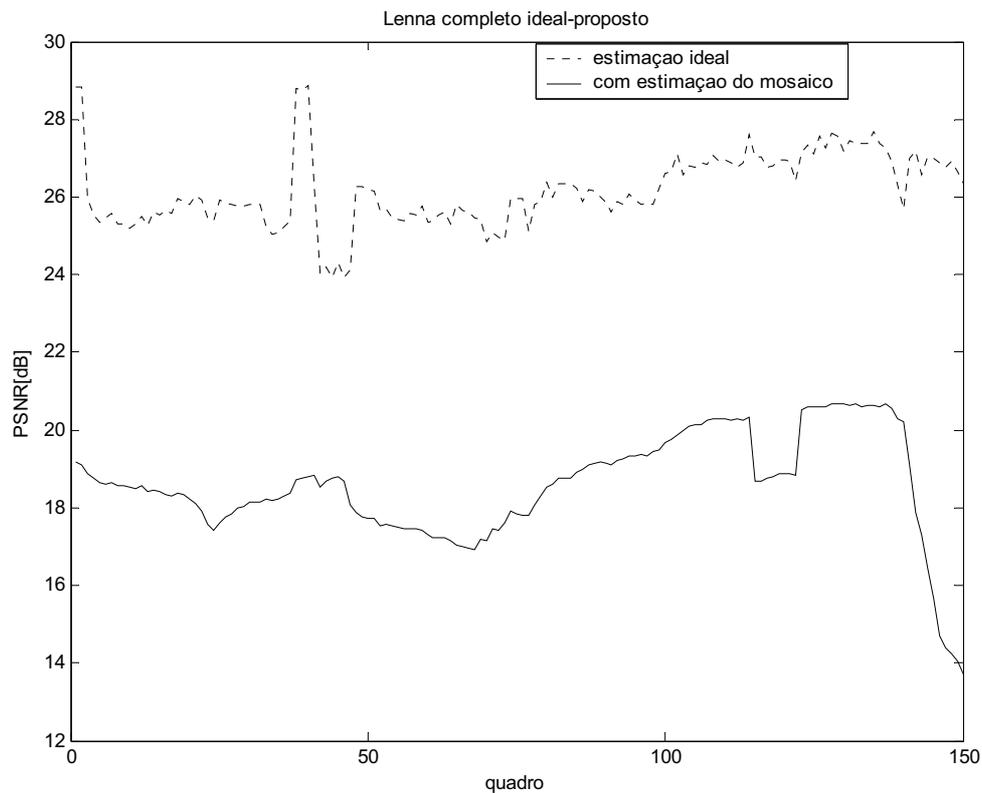
No método de estimação proposto neste trabalho o erro de reconstrução decorrente da interpolação pode ser considerado como um limite máximo, em

termos de qualidade de reconstrução, que um método pode atingir. Portanto, os gráficos dos PSNR estão estreitamente relacionados à interpolação.

Através dos parâmetros de movimento global dos vídeos artificiais, que já são conhecidos, foram feitos testes para ver a influência da interpolação na qualidade da reconstrução. Os testes foram realizados com as oito seqüências artificiais, que serão apresentadas a seguir.

#### **4.4.1. Análise dos resultados**

Nos resultados da seqüência Lenna\_completo, apresentados na Figura 49, é possível ver uma certa relação entre o PSNR com os parâmetros ideais e o PSNR com os parâmetros estimados. Alguns máximos locais e mínimos locais estão localizados no mesmo quadro. Por exemplo, o máximo local próximo ao quadro 40 está presente nos dois gráficos. Outro exemplo é o mínimo local que está próximo ao quadro 20, que também está presente nos dois gráficos. Porém, através deste gráfico é possível observar que estimativa tem uma diferença de qualidade em relação ao ideal em média por 17dB.

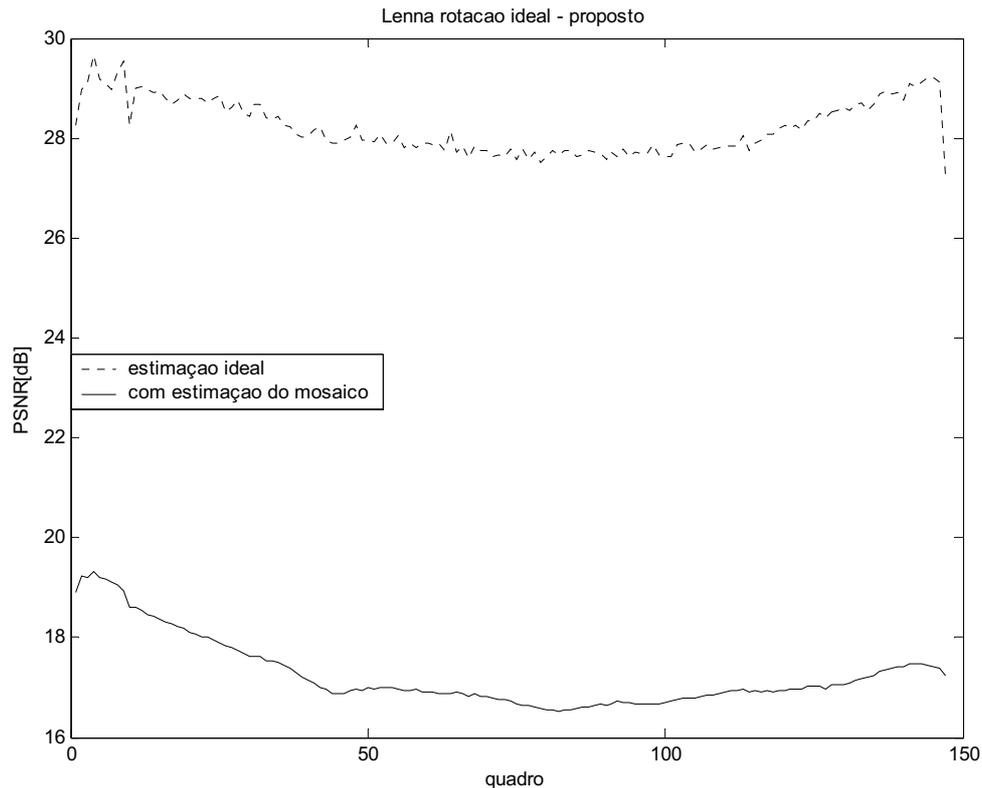


**Figura 49-** Relação entre a interpolação e o PSNR, Lenna\_completa.

De um modo geral, as seqüências de rotação possuem o PSNR da estimação com um comportamento global semelhante ao PSNR dos parâmetros ideais. Portanto, é percebido que mesmo em situações em que os parâmetros de movimento global não são tão bem estimados, ocorre uma influência da interpolação. O comportamento do PSNR, apesar de ter diferenças muito grandes de valores, apresenta alguns mínimos e máximos locais nos mesmos quadros. Tendo a interpolação uma influência grande no comportamento dinâmico do PSNR, variações em torno de um valor médio, para estes três casos. Dessa forma é possível concluir que, em algumas situações, a estimação de movimento global pouco pode fazer para corrigir a qualidade da reconstrução dos quadros a partir do mosaico.

Na seqüência Lenna\_rotação, Figura 50, é perceptível que a interpolação possui uma influência nas variações do gráfico do PSNR. Como no caso da região

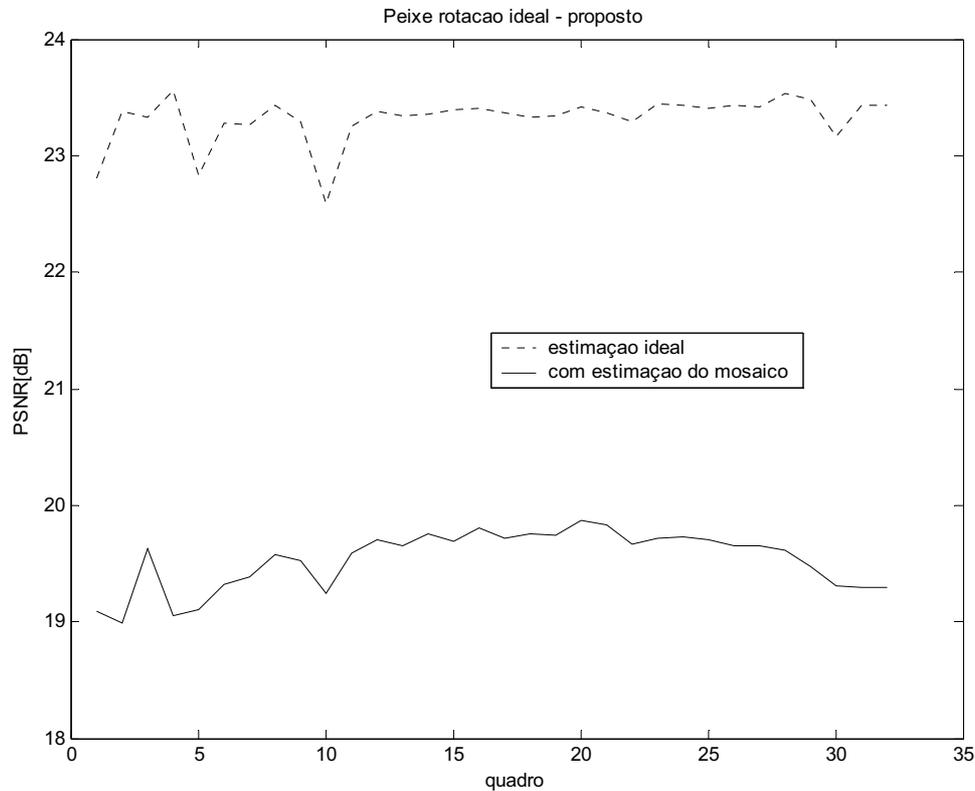
dos quadros iniciais da seqüência de vídeo, em que ocorrem os máximos globais de qualidade. Porém, também é perceptível que a influência da interpolação é atenuada pela má estimativa dos parâmetros de movimento global. Pois, os valores de PSNR estão muito distantes, em aproximadamente 10dB de diferença média.



**Figura 50-** Relação entre a interpolação e o PSNR, Lenna\_rotação.

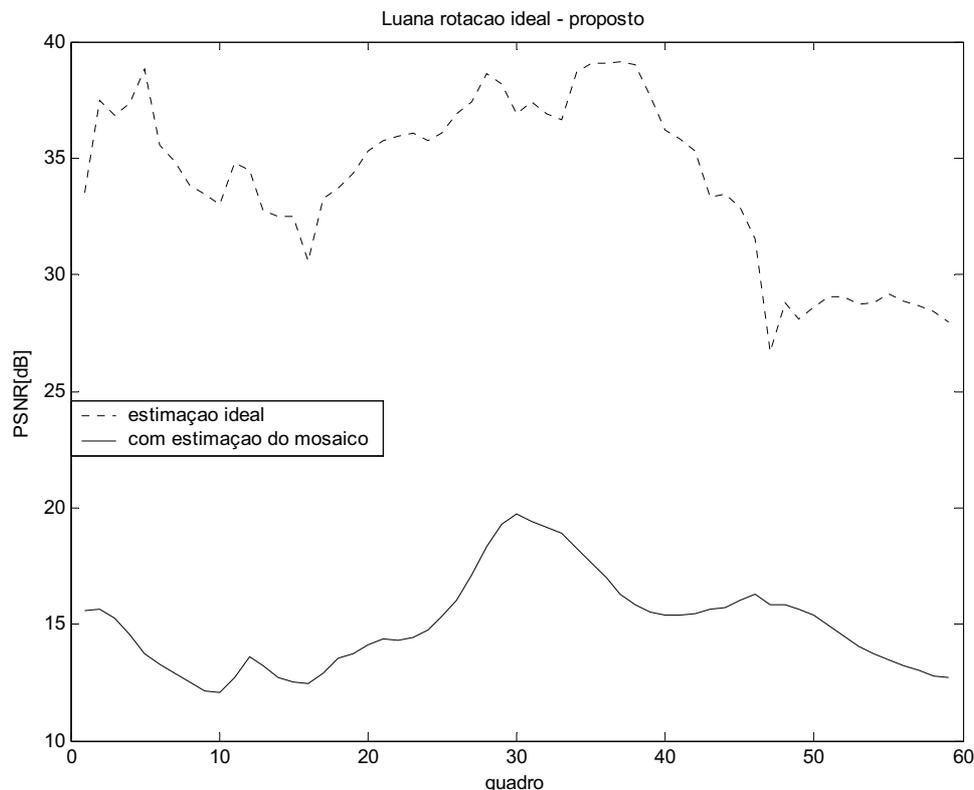
Em peixe\_rotação, Figura 51, também é perceptível que a interpolação possui uma influência nas variações do gráfico do PSNR. Mas neste caso, é possível ver uma relação mais direta do que no caso anterior. Pois os parâmetros de movimento global foram mais bem estimados. Neste caso vemos vários máximos locais e mínimos locais que coincidem, não em valores, mas na posição do quadro. Outra questão é a diferença entre os gráficos de PSNR, que possuem uma diferença média de 4dB, bem menor que a anterior. Esta melhora, na

qualidade em relação à interpolação, está relacionada à seqüência peixe\_rotação ter um menor número de quadros do que a seqüência Lenna\_rotação.



**Figura 51-** Relação entre a interpolação e o PSNR, peixe\_rotação.

Luana\_rotação, Figura 52, é uma seqüência em que a influência da interpolação é extremamente atenuada, se comparada com as seqüências anteriores. Isto pode ser verificado pelo comportamento global dos PSNR. O PSNR ideal possui suas variações muito mais abruptas do que o PSNR estimado. Porém, o fator que mais diferencia as duas curvas é a diferença média de valores que está por volta de 28dB. Portanto, os erros estão mais relacionados, neste caso, a má estimativa dos parâmetros de movimento global do que à interpolação.

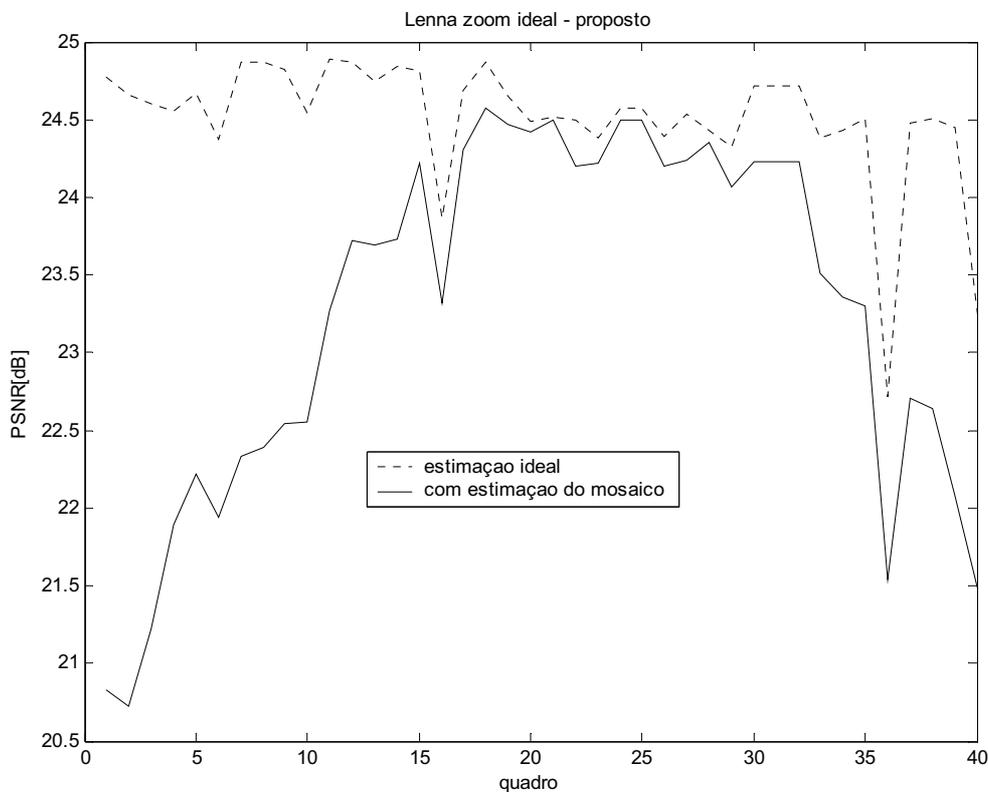


**Figura 52-** Relação entre a interpolação e o PSNR, Luana\_rotação.

As três seqüências de zoom possuem um comportamento mais influenciado pela interpolação do que nos casos anteriores. Em alguns trechos, de alguns gráficos, o PSNR estimado se confunde, em comportamento dinâmico e até mesmo numericamente, com o PSNR ideal. Nestes casos, é possível identificar não somente uma influência no comportamento dinâmico do PSNR, mas também uma influência da interpolação nos valores numéricos dos PSNR. Esta maior influência da interpolação advém do fato da estimação dos parâmetros de movimento global serem mais bem estimados nas situações de zoom. A interpolação, nos três casos de zoom, é transformada em um limite de qualidade, que quando atingido, torna qualquer trabalho de refinamento da qualidade do mosaico, além do PSNR ideal, inútil.

Na seqüência Lenna\_zoom, Figura 53, é mais visível, a influência da interpolação, no trecho entre os quadros 15 e 35. Neste trecho, o PSNR da

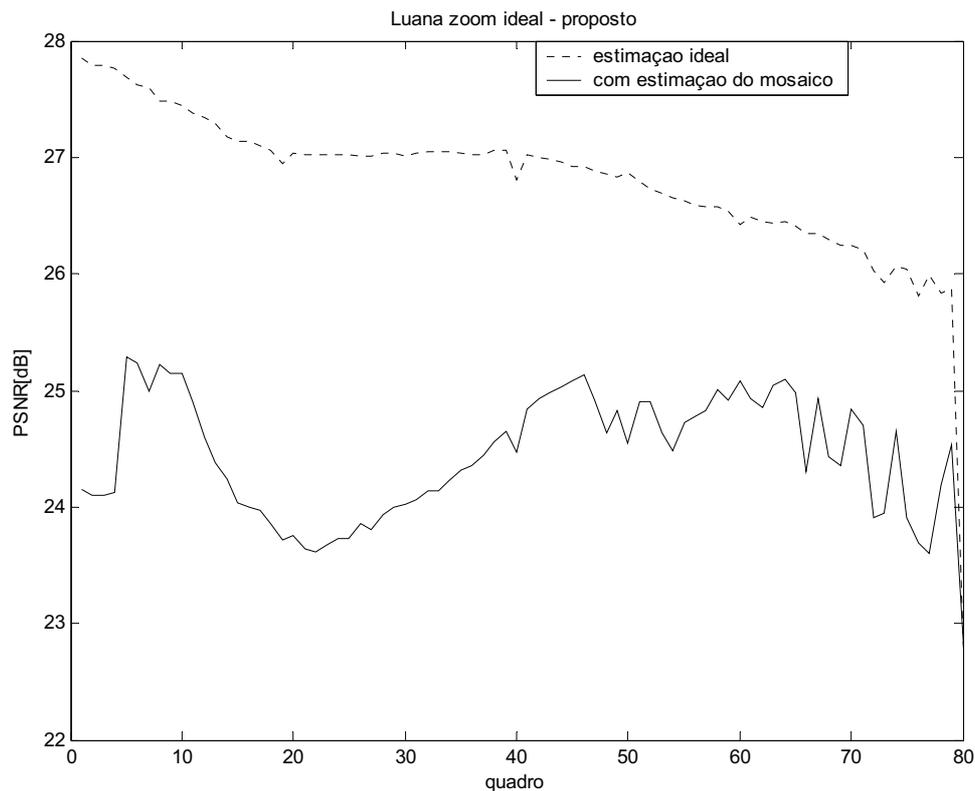
estimação com mosaico têm mínimos com posições coincidentes, não numericamente iguais, mas com valores bem próximos. Como no quadro 16, com 0,6dB de diferença, e no quadro 36, com 1,5dB de diferença. Em outros pontos deste intervalo, como entre os quadros 21 e 25, as diferenças são menores do que 1dB. Porém, no comportamento global vemos uma atenuação maior do PSNR estimado pelo mosaico em alguns trechos, como o trecho inicial com 3,5dB de diferença e no trecho final com 3dB de diferença.



**Figura 53-** Relação entre a interpolação e o PSNR, Lenna\_zoom.

Das três seqüências de zoom Luana\_zoom, Figura 54, é a que apresenta uma relação menos estreita, em termos da qualidade de reconstrução dos quadros a partir do mosaico, com a interpolação. É observável que nos primeiros quadros, o PSNR estimado através do mosaico não possui um comportamento muito semelhante, devido à estimação dos parâmetros de movimento global.

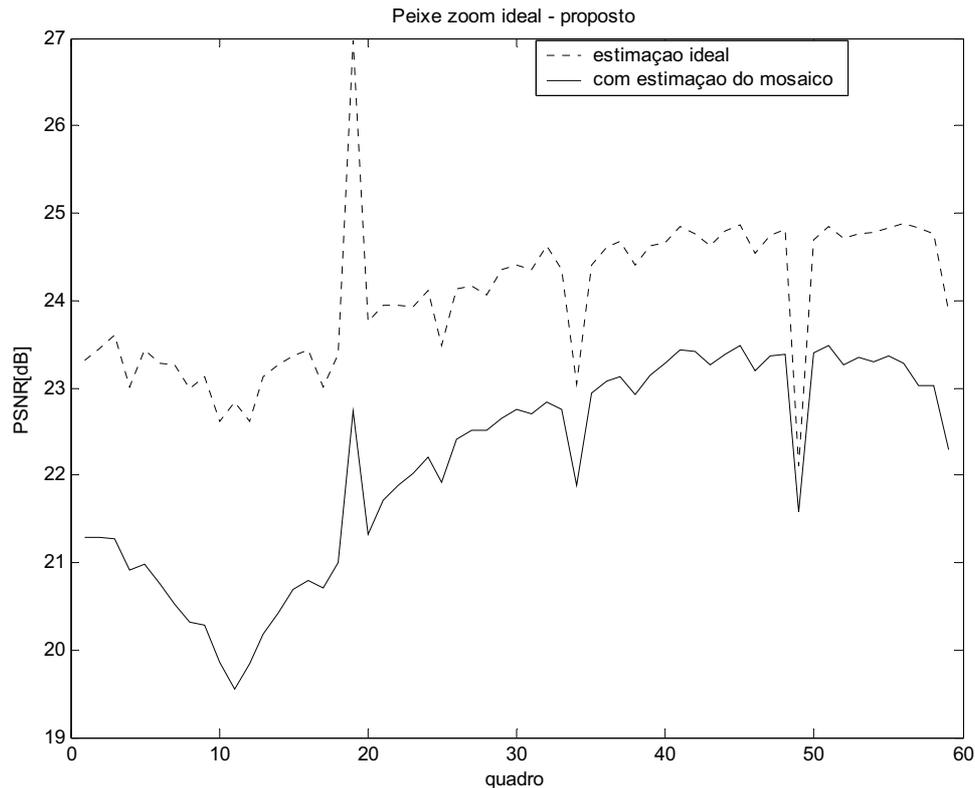
Porém em outros trechos, mais no final da seqüência de vídeo, existe uma proximidade maior dos valores do PSNR com estimação do mosaico. Esta aproximação, nos trechos finais está relacionada a maior quantidade de texturas, zoom maior com menos regiões homogêneas, que melhoram a qualidade do campo de vetores.



**Figura 54-** Relação entre a interpolação e o PSNR, Luana\_zoom.

A exemplo da seqüência Lenna\_zoom, a seqüência peixe\_zoom, Figura 55, também apresenta pontos em que os valores dos PSNR muito próximos. Das três seqüências de zoom é a que possui uma relação mais estreita, visualmente, entre o PSNR estimado do mosaico e o PSNR da estimação ideal. O quadro em que os valores estão mais próximos é o 49, com menos de 0,5dB de diferença. Além de ter um comportamento médio bastante próximo ao PSNR da estimação ideal, com diferença média de aproximadamente 2dB. Além da proximidade dos valores é

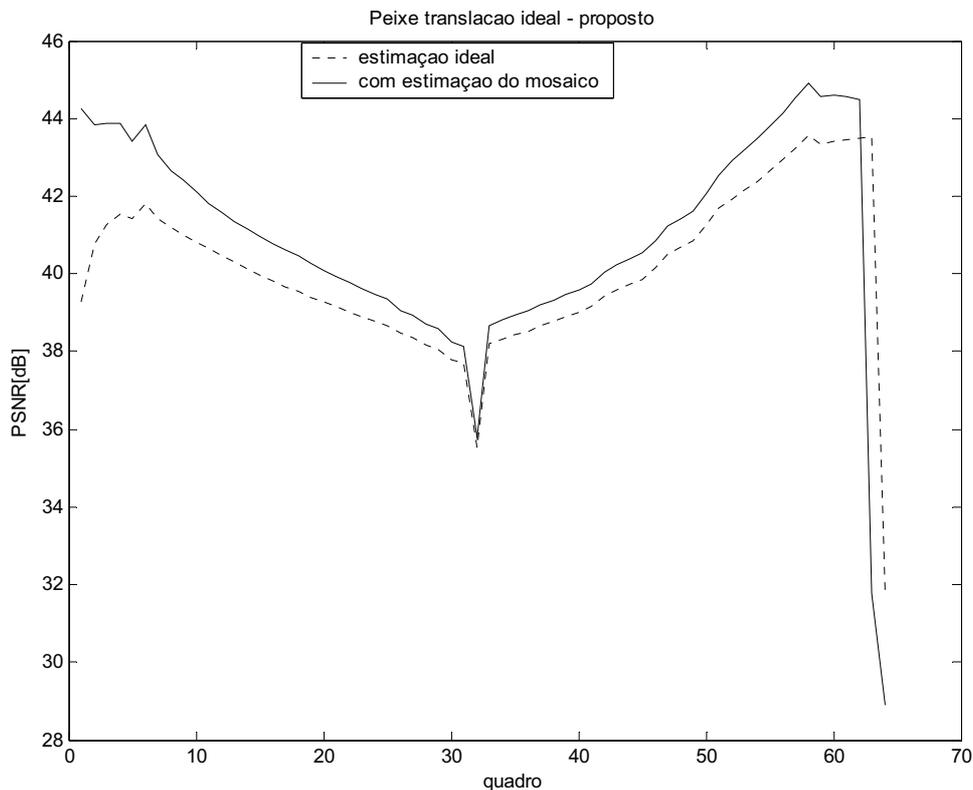
visível que o comportamento dinâmico dos valores dos dois PSNR é bastante similar. Nesta seqüência, é possível dizer que os erros de reconstrução são muito mais relacionados à interpolação do que a estimação dos parâmetros de movimento.



**Figura 55-** Relação entre a interpolação e o PSNR, peixe\_zoom.

A translação é o caso em que o método está mais adaptado. Portanto, é também a situação em que a interpolação terá maior influência nos erros de reconstrução. No gráfico abaixo, é visível que os erros de reconstrução estão intimamente ligados aos erros decorrentes da interpolação. Pois, a diferença média entre os dois PSNR está em 2dB, como no caso do peixe\_rotacao. Em praticamente todos os trechos da seqüência de vídeo, a estimação por mosaico possui valores que seguem a tendência da estimação ideal. Nesta seqüência os erros de reconstrução não decorrem de erros na estimação dos parâmetros de

movimento global. Mas sim de erros de interpolação no momento da construção do mosaico e da reconstrução dos quadros a partir do mosaico. Estes erros ocorrem quando as translações não são inteiras e existe a necessidade de interpolar um pixel fracionário.



**Figura 56-** Relação entre a interpolação e o PSNR, peixe\_translação.

Através destes resultados foi possível verificar que em alguns casos, como translação e zoom, as estimações com mosaico estão perto do limite imposto pela interpolação. E na seqüência completa e nas rotações, os erros estão mais relacionados com os erros de estimação dos parâmetros de movimento global. Porém, mesmo em seqüências com erros muito grandes, devido à estimação de parâmetros, existe influência dos erros de interpolação. Portanto, é possível concluir que a interpolação realmente tem um papel fundamental na construção do mosaico e na reconstrução dos quadros. Sendo impossível atingir erros nulos de reconstrução, neste método de estimação de mosaico, mesmo com uma

estimação dos parâmetros de movimento global ideal. Uma forma de atenuar estes erros seria a utilização de um algoritmo de interpolação mais poderoso, porém isto aumentaria o custo computacional do processo.

#### **4.5. Avaliação do Método de Estimação dos Parâmetros de Movimento Global**

Nesta seção será avaliado o método de estimação dos parâmetros de movimento global, sem levar em conta a qualidade do mosaico resultante do processo. Para a avaliação deste método foram criados vários vídeos artificiais a partir de imagens panorâmicas. Os parâmetros das transformações geométricas, de cada quadro em relação à imagem panorâmica, são conhecidos. O método de estimação de movimento global será avaliado em função dos parâmetros de movimento global conhecidos das seqüências de vídeo. Serão analisados, nesta seção o desempenho do método em relação à translação, ao zoom, a rotação e a todas as transformações juntas, como nas seções anteriores. Serão analisados somente os parâmetros que possuem alguma influência efetiva nas seqüências.

#### 4.5.1. Análise dos resultados

Na seqüência Lenna\_completo o desempenho para os parâmetros de zoom, Figura 57, foi satisfatório. Até o quadro 45, o método acompanhou bem os parâmetros ideais. A partir do quadro 46, os parâmetros de zoom estimados começam a se distanciar dos parâmetros ideais. Isto decorre, muito provavelmente, de um erro de estimação que esta presente nos parâmetros de rotação no quadro 27. Porém, o método corrige o erro, aos poucos, e encontra os parâmetros ideais no quadro 130. De um modo geral os parâmetros de zoom, foram bem modelados pela estimação. Porém, alguns erros não foram evitados, como a influência da rotação na estimação dos parâmetros de zoom. No caso do parâmetro  $a_3$  o estimado acompanha em praticamente todos os quadros o parâmetro ideal.

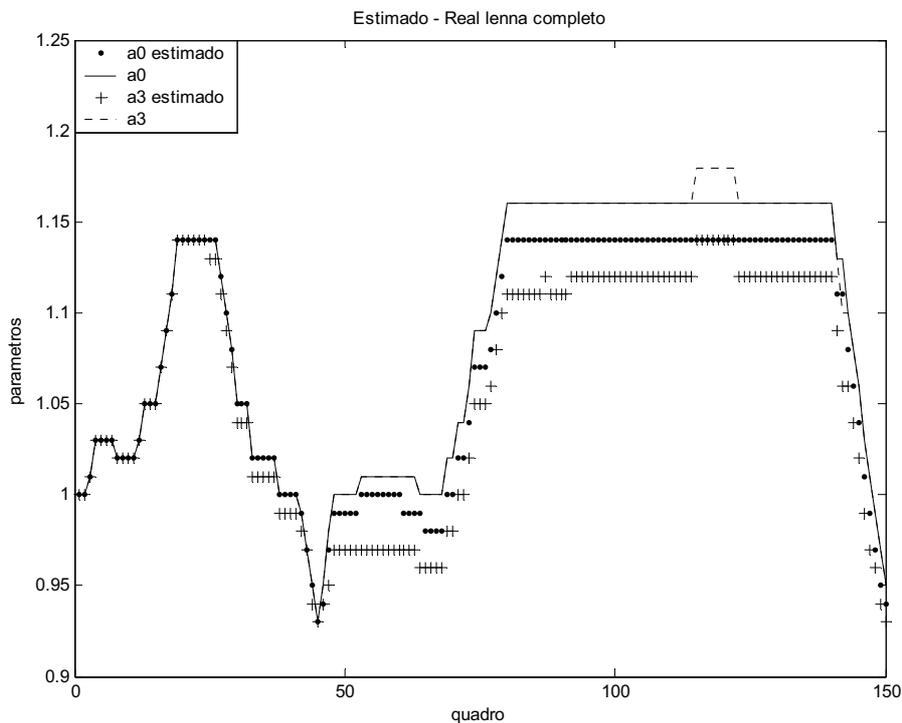
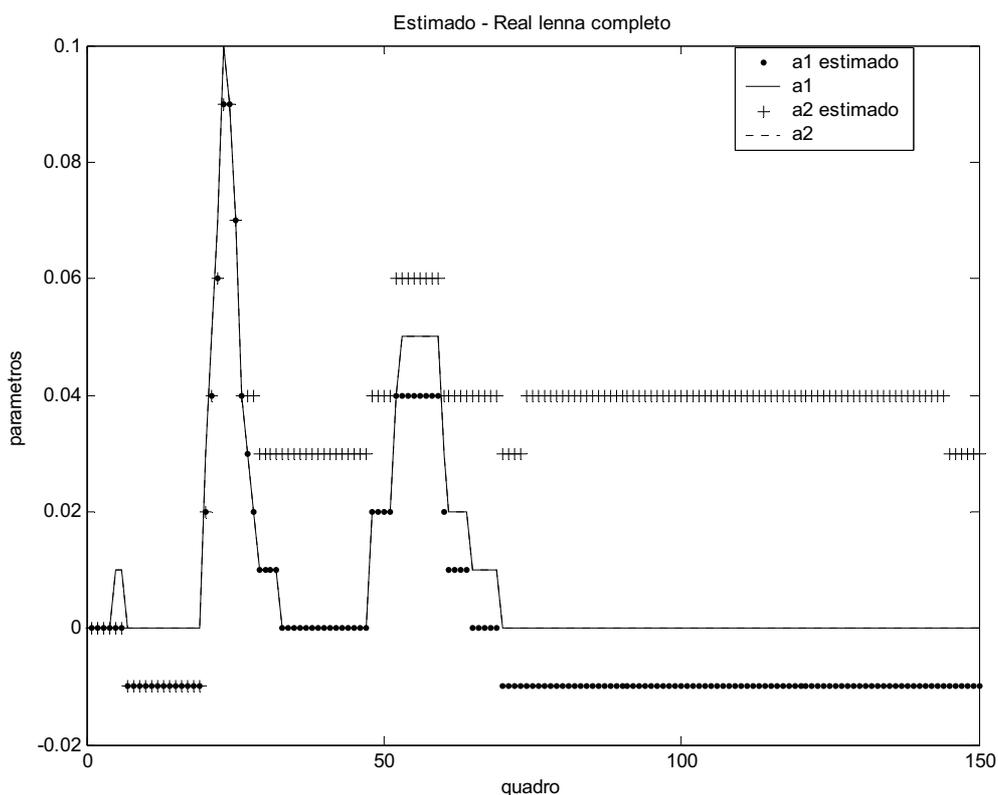


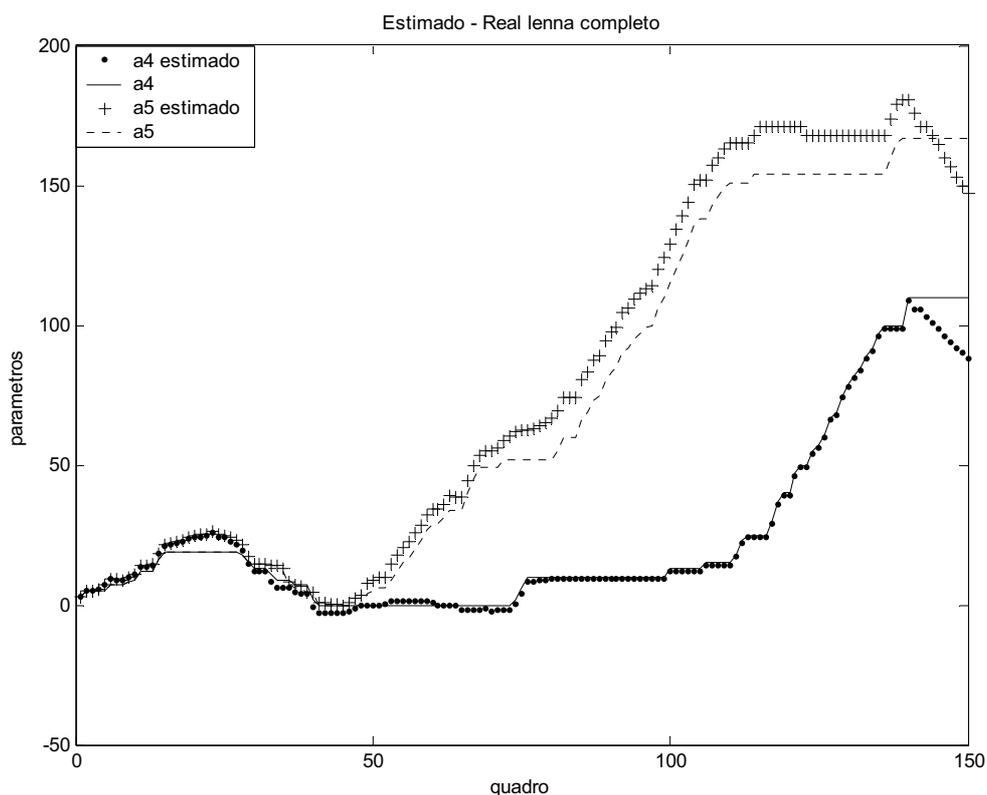
Figura 57- Relação entre os parâmetros de movimento estimados e os reais I, Lenna\_completo.

Na Figura 58 são apresentados os parâmetros de rotação, neste caso o parâmetro  $a_2$  ideal possui valores idênticos ao parâmetro  $a_1$  ideal em todos os quadros. Os parâmetros de rotação são pontos críticos do método de estimação, ocorrendo alguns problemas para acompanhar os parâmetros de rotação ideais. Nos primeiros quadros, no quadro 6, ocorre um erro de estimação. Porém, o método corrige o erro entre os quadros 16 e 29. Do quadro 29 em diante o erro de estimação não é mais corrigido. Um grande problema do erro na estimação dos parâmetros de rotação é a propagação desse erro nos outros parâmetros, principalmente nos parâmetros de zoom. Esta propagação de erros ocorre, devido à tentativa do método de corrigir, nos outros parâmetros, o erro causado pelos parâmetros de rotação.



**Figura 58-** Relação entre os parâmetros de movimento estimados e os reais II, Lenna\_completo.

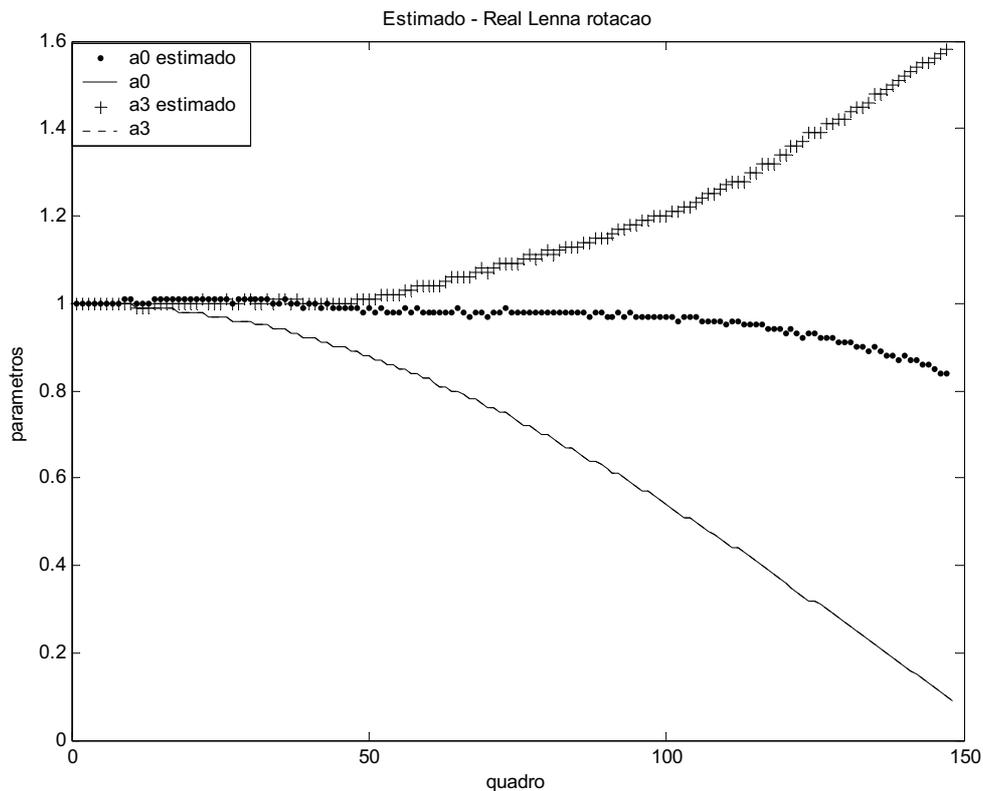
Os parâmetros de translação, Figura 59, foram os que tiveram melhor desempenho, dentre todos os outros. Como os parâmetros de translação são menos sensíveis à propagação dos erros, os efeitos dos erros na estimação dos parâmetros de rotação foram menos sentidos. Os parâmetros de translação estimados seguem a tendência dos parâmetros de translação ideais, em praticamente todos os trechos da seqüência. Porém, em alguns pontos a diferença entre os parâmetros de translação estimados e os parâmetros de translação ideais foi grande, 16 pixels entre o quadro 114 e 122. Que percentualmente equivale a um erro de cerca de 10%. Esses erros na estimação dos parâmetros de translação são, muito provavelmente, acarretados pelos erros na estimação dos parâmetros de rotação.



**Figura 59-** Relação entre os parâmetros de movimento estimados e os reais III, Lenna\_completo.

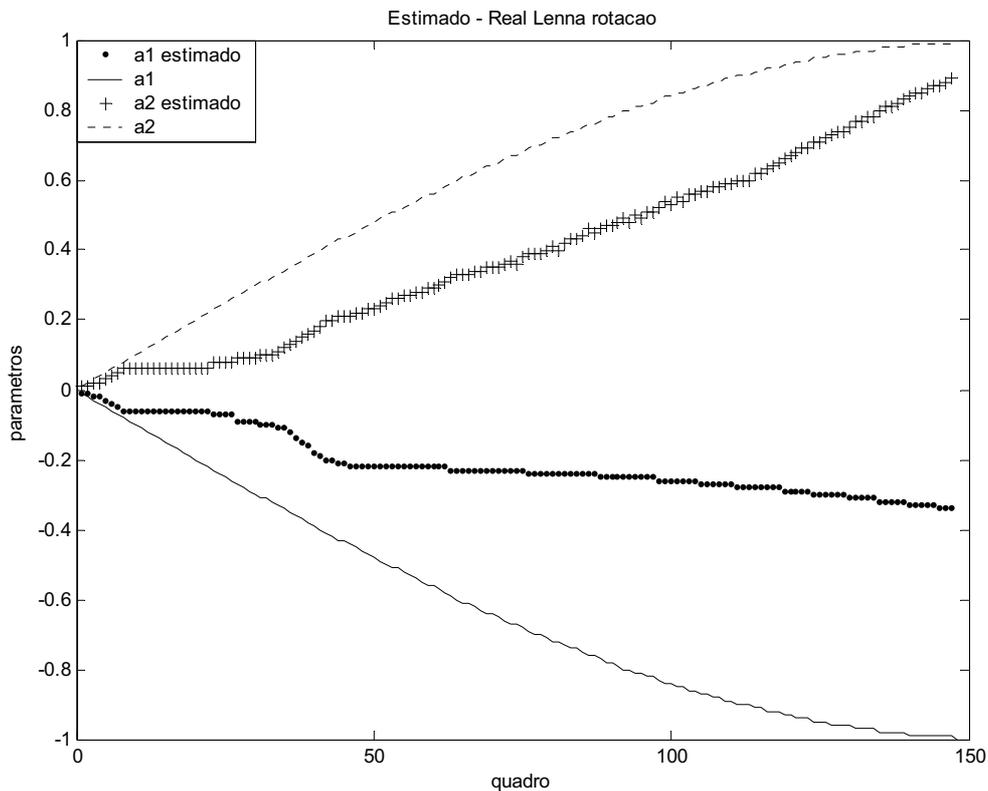
Em Lenna\_rotação somente os parâmetros que influenciam em uma rotação pura seriam necessários, os dois parâmetros de zoom e os dois parâmetros de rotação, para a estimação. Porém, devido aos erros de estimação os parâmetros de translação também se manifestaram.

O método de estimação dos parâmetros de movimento global, para os parâmetros de zoom, Figura 60, não teve um desempenho muito satisfatório. Neste caso o parâmetro  $a_0$  ideal possui valores idênticos ao parâmetro  $a_3$  ideal em todos os quadros. Somente no início da seqüência os parâmetros estimados conseguiram acompanhar os parâmetros ideais. No restante da seqüência os parâmetros ideais e os parâmetros estimados chegaram a ter tendências divergentes. Este problema na modelagem dos parâmetros de zoom, muito provavelmente, é decorrente da excessiva classificação dos vetores de movimento. Porém outros fatores, como campos de vetores de movimento mal estimados, influenciaram na divergência dos parâmetros de zoom também.



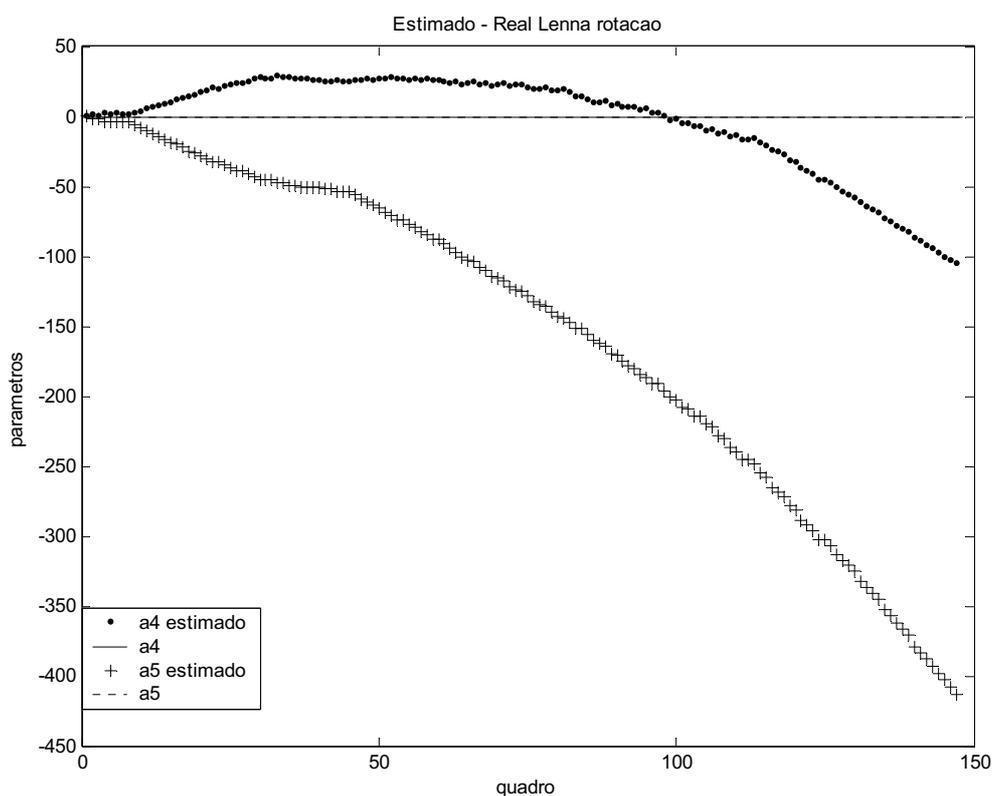
**Figura 60-** Relação entre os parâmetros de movimento estimados e os reais I, Lenna\_rotação.

Os parâmetros de rotação estimados, Figura 61, tiveram uma estimação melhor, nesta seqüência, do que os parâmetros de zoom. Apesar dos parâmetros de rotação estimados estarem com erros grandes, eles ainda seguem a tendência dos parâmetros de zoom ideais. Inicialmente os parâmetros de rotação estimados estão muito próximos dos parâmetros ideais, principalmente nos primeiros dez quadros. Porém, a partir do quadro 10, vários erros de estimação começam a ocorrer. Dentre estes vários erros podem ser citados os problemas na coerência do campo de vetores de movimento, problemas na classificação dos vetores movimento e os acúmulos de pequenos erros.



**Figura 61-** Relação entre os parâmetros de movimento estimados e os reais II, Lenna\_roteação.

Os parâmetros de translação, Figura 62, refletem bem os erros de estimação. Se a estimação das rotações estivesse coerente, os parâmetros de translação não teriam uma presença muito marcante. Devido aos erros de estimação já citados nos gráficos anteriores, os parâmetros de translação apresentam valores de estimação bem diferentes dos parâmetros ideais. Os parâmetros de translação estimados, devido ao algoritmo, tendem a corrigir os erros, na estimação do mosaico, ocasionados pelos outros parâmetros de movimento. Porém, esta compensação no decorrer do processo torna a estimação ainda pior. O que gera, em alguns quadros, grandes erros de estimação, principalmente no final da seqüência.



**Figura 62-** Relação entre os parâmetros de movimento estimados e os reais III, Lenna\_roteação.

As estimações dos parâmetros de movimento global das seqüências peixe\_rotação e Luana\_rotação possuem os mesmos problemas da seqüência Lenna\_rotação.

Fica visível, pela comparação entre os parâmetros de zoom estimados e os parâmetros de zoom ideais, que os problemas de estimação são graves. Pois, somente nos primeiros quadros não temos erros de estimação. Estes erros de estimação são decorrentes, como no caso do Lenna\_rotação, de problemas no campo de vetores de movimento e da classificação dos vetores de movimento. Os pequenos erros em quadros específicos geram um acúmulo de erros, que atrapalham todo o processo de estimação.

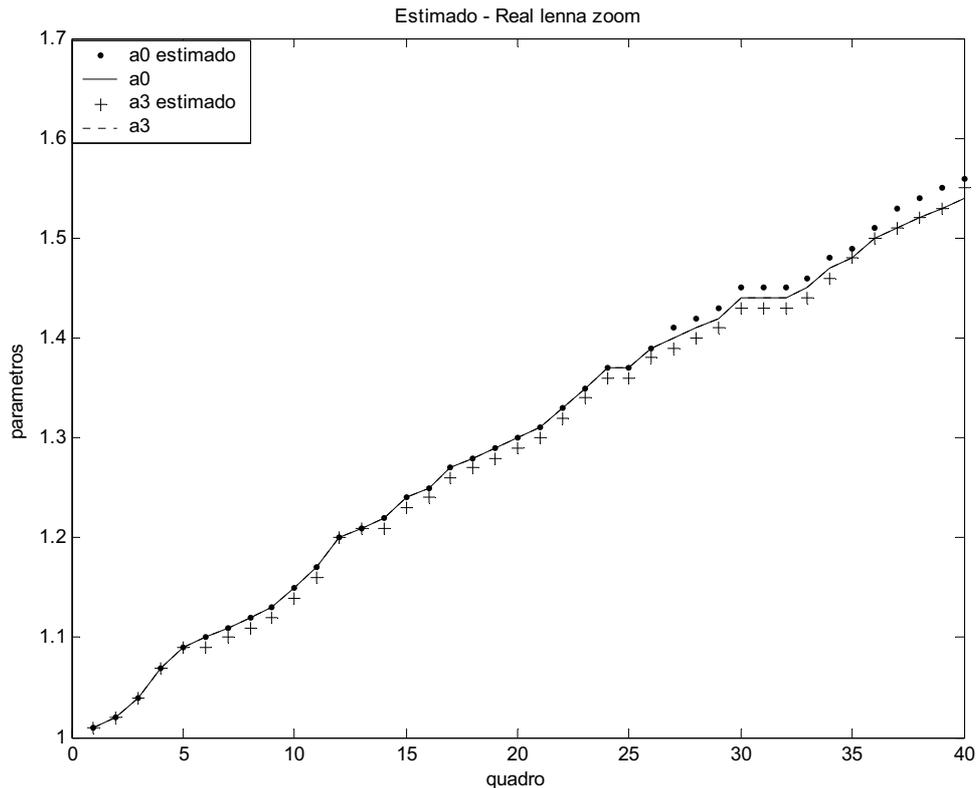
De maneira geral, a estimação dos parâmetros de movimento global para as seqüências de rotação não apresentou um comportamento satisfatório. Graças aos erros no processo de estimação, erros de estimação de campo de vetores de movimento, erros na classificação dos vetores e o acúmulo de erros, o método não obteve o êxito previsto. Uma forma de melhorar o desempenho do método seria a modificação de alguns limiares presentes no método. Porém, seria necessário verificar o que as modificações de limiares trariam de malefícios para as outras situações de movimento global, como translação e zoom.

As seqüências de zoom possuem, de forma geral, estimções de movimento global, com este método, melhores do que os casos de rotações. São três as razões principais para a facilidade na estimação de zoom. Primeiro, o zoom necessita de um menor número de parâmetros para ser modelado. Segundo, a utilização de parâmetros com menor sensibilidade do que o caso da rotação. Terceiro, o método de estimação, empregado nas seqüências de vídeo, está muito mais adaptado às situações de zoom.

Na seqüência Lenna\_zoom os parâmetros de rotação não se manifestaram. Isto ocorreu devido à saturação, em zero, dos parâmetros de rotação.

A seqüência Lenna\_zoom possui um resultado satisfatório para os parâmetros de zoom, Figura 63. É percebido, através do gráfico, que os parâmetros de zoom estimados seguem, em toda seqüência, a tendência dos parâmetros ideais de zoom. Para demonstrar como os erros foram pequenos entre

os parâmetros de zoom ideais e os parâmetros de zoom estimados, podem ser mostradas as diferenças percentuais. A diferença percentual média entre os parâmetros de zoom estimados e ideais é de 0,3% para  $a_0$  e 0,5% para  $a_3$ . E a diferença percentual máxima é de 1,3%, no quadro 37.

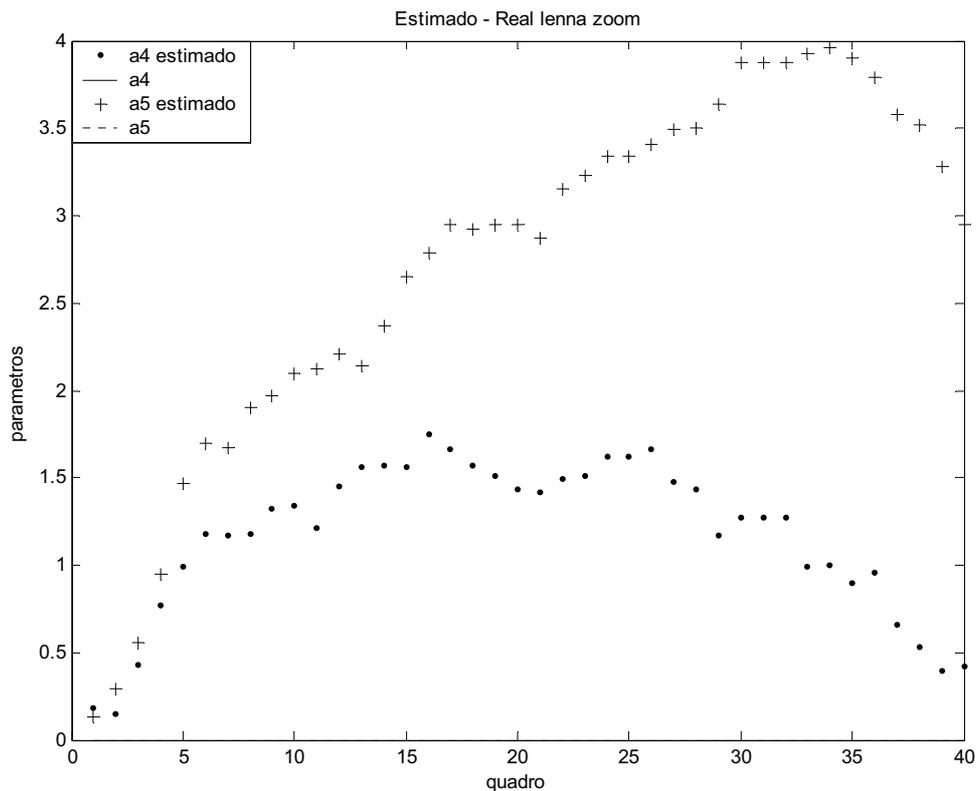


**Figura 63-** Relação entre os parâmetros de movimento estimados e os reais I, Lenna\_zoom.

Como a estimação dos parâmetros de zoom não é perfeita, os parâmetros de translação se manifestam, Figura 64, pois neste caso os parâmetros de translação possuem valores nulos em toda a seqüência. O comportamento dos parâmetros de translação retrata, até de uma forma amplificada, os erros na estimação dos parâmetros de zoom. A região em que o erro de estimação nos parâmetros de zoom está mais evidente é entre os quadros 30 e 40. Que também é a região em que os parâmetros de translação apresentam seus maiores valores,

se somados. Portanto, nesse caso, os erros dos parâmetros de translação estão intimamente ligados aos erros dos parâmetros de zoom.

Se os erros dos parâmetros de translação, nesta seqüência, forem comparados com erros das seqüências de rotação, é possível perceber que são erros pequenos. Porém, mesmo pequenos, estes erros influenciarão no resultado final do processo.

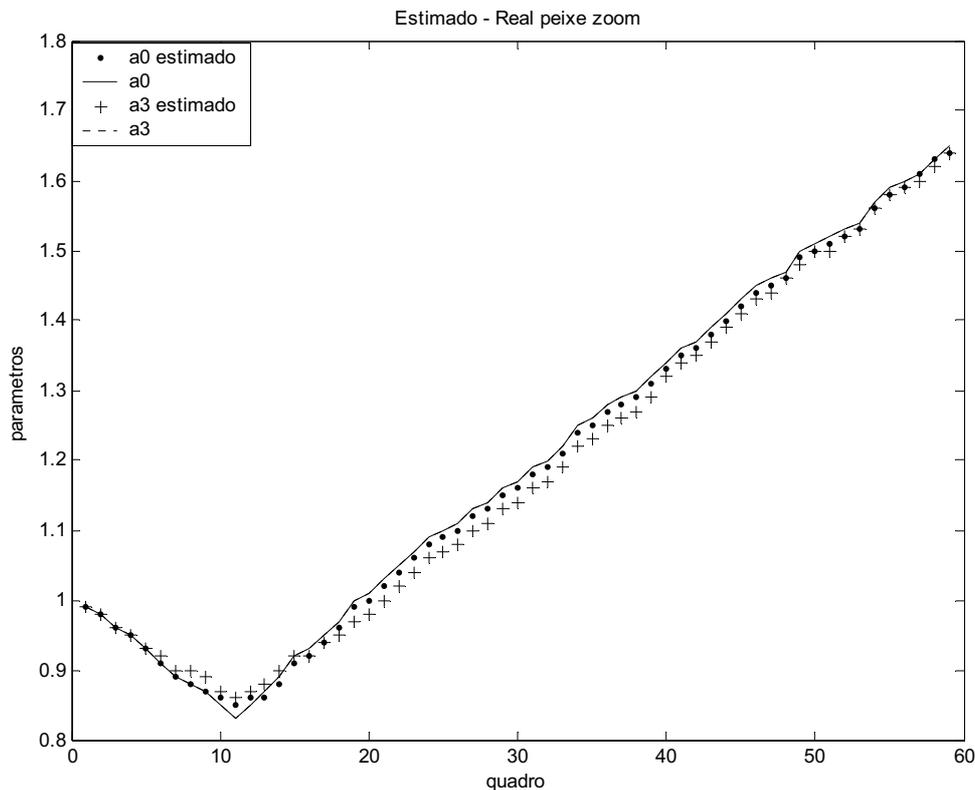


**Figura 64-** Relação entre os parâmetros de movimento estimados e os reais II, Lenna\_zoom.

Como na situação anterior, os parâmetros de rotação não se manifestaram, na seqüência peixe\_zoom.

A estimação dos parâmetros de zoom, Figura 65, também apresentou um desempenho satisfatório, na seqüência peixe\_zoom. Os parâmetros de zoom estimados acompanharam, em toda a seqüência, a tendência dos parâmetros ideais. Sendo que nesta seqüência ocorreu zoom para fora, *zoom out*, e zoom para dentro, *zoom in*, ao invés da situação anterior em que ocorreu somente zoom

para fora. Assim, foi percebido que mesmo em situações em que existe mudança de sentido no zoom o método consegue acompanhar. Em termos de erro percentual, neste caso, o erro máximo foi de 3,6%, no quadro 11, e o erro médio foi de 0,7% para  $a_0$  e 1,6% para  $a_3$ .

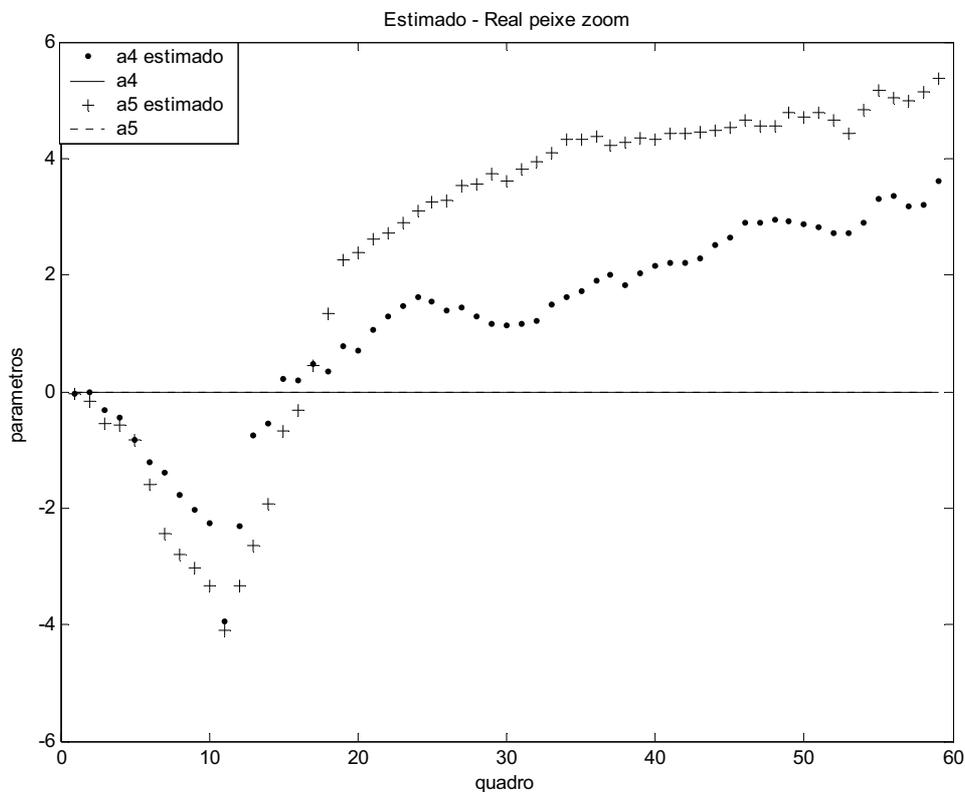


**Figura 65-** Relação entre os parâmetros de movimento estimados e os reais I, peixe\_zoom.

Como em Lenna\_zoom, é possível ver nesta seqüência que os parâmetros de translação, Figura 66, retratam o erro de estimação dos parâmetros de zoom. Nesta seqüência é possível perceber que o *zoom in* é confundido com uma translação no sentido negativo, em alguns pontos. E que o *zoom out* é confundido com uma translação no sentido positivo. Esta confusão entre a estimação dos parâmetros de zoom e translação, deve estar relacionada à distribuição irregular dos vetores de movimento. Como o método é baseado na filtragem dos vetores de movimento, por muitas vezes pode ocorrer uma distribuição irregular de vetores de movimento, em torno do centro de zoom. A distribuição irregular pode privilegiar

um lado em torno do centro de zoom em detrimento de outro. Dessa forma, o campo de vetores filtrados pode retratar um zoom com uma certa translação, ao invés de um zoom puro.

Em termos de erros, nos parâmetros de translação desta seqüência, podem ser considerados razoáveis. Porém, em comparação com a seqüência anterior os erros são maiores. Isso se deve, muito provavelmente, a mudança de sentido do zoom, que danifica um pouco a estimação dos parâmetros.



**Figura 66-** Relação entre os parâmetros de movimento estimados e os reais II, peixe\_zoom.

Os resultados da seqüência Luana\_zoom têm um comportamento bastante semelhante aos dois casos anteriores, principalmente ao primeiro caso. Portanto não será apresentado nesta seção.

De modo geral, é possível dizer que a estimação dos parâmetros de movimento, em seqüências de zoom, é satisfatório. Também é possível concluir,

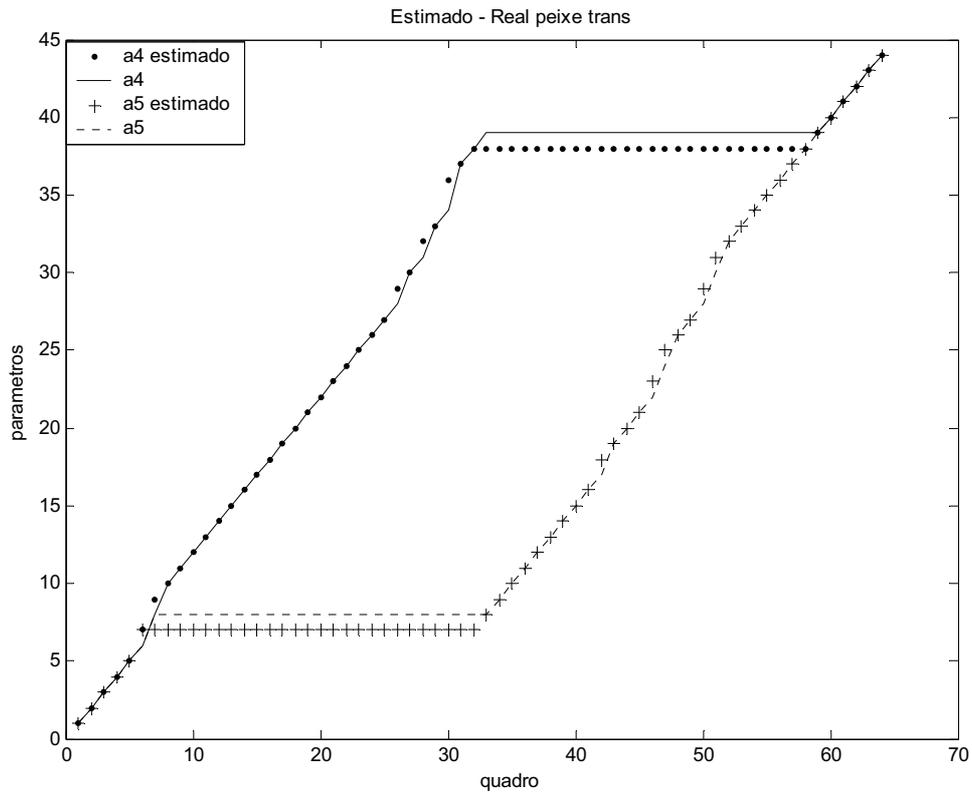
que os parâmetros de translação estão relacionados aos erros de estimação nos parâmetros de zoom. Outra conclusão, é que o método está muito mais apto a estimar os parâmetros de movimento de seqüências de zoom do que seqüências de rotação.

A translação é a situação de melhor desempenho do método. Em seqüências de pura translação, além dos pequenos erros de estimação obtidos, não é observada a interferência entre as estimações dos parâmetros. Isto se deve, em muito, ao método que trata de forma diferente as translações puras.

Na seqüência *peixe\_translação* os parâmetros de translação, Figura 67, seguem em toda a seqüência os parâmetros de translação ideais. E também é observável que os parâmetros de rotação e de zoom não interferem nas estimações, pois todos são nulos. O erro entre os parâmetros estimados e os parâmetros ideais não ultrapassa um *pixel*. O erro percentual médio, entre os parâmetros estimados e os parâmetros ideais, é de 0,028% para a translação em *x* e de 0,017% para a translação em *y*.

Os erros de estimação ocorrem em função de mudanças abruptas nos parâmetros de translação em *x* e em *y*. Os erros se propagam durante aproximadamente 20 quadros. Porém, o método corrige os erros, a partir de outras duas mudanças abruptas.

De modo geral o erro entre os parâmetros é baixo, se comparado com as estimações de seqüências de zoom e de rotação.



**Figura 67-** Relação entre os parâmetros de movimento estimados e os reais I, peixe\_translação.

A estimação dos parâmetros de movimento global, para seqüências de translação, é praticamente perfeita. Isto se deve, além do mérito do método, aos vetores de movimento, que nestas situações são mais coerentes. Outra questão é a modelagem do movimento de translação, que não depende de termos multiplicativos, que são bem mais sensíveis a variações.

## 4.6. Comparação com outros métodos

Esta seção será um comparativo com outros métodos de estimação de mosaico e também com outras formas de compensação de movimento. Para comparar os resultados de outros trabalhos serão utilizados os valores de PSNR da reconstrução dos quadros a partir do mosaico, nos casos de compensação por mosaico. E nos casos de compensações quadro a quadro, será utilizada a reconstrução do quadro posterior através do quadro anterior. Nas compensações quadro a quadro serão utilizadas, para comparação, a compensação clássica por busca de blocos, a compensação utilizada na seção 4.2, que utiliza o método proposto para estimação de movimento global, e a diferença pura entre os quadros consecutivos. Os testes utilizam o PSNR médio de 150 quadros para fazer este comparativo:

**Tabela 1-** Comparação entre os métodos de estimação de mosaico .

Métodos	Coast_guard PSNR [dB]	Stefan PSNR [dB]
Método Proposto, mosaico.	17,10	16,85
Método Proposto, compensação quadro a quadro.	26,78	24,74
Diferença quadro a quadro.	23,78	19,11
Compensação quadro a quadro, clássica, por semelhança de blocos.	30,29	27,42
S. Fukunaga [25], estimação feita diretamente do mosaico.	23,58	20,88
A. Smolic [8].	20,23	20,34
S. Fukunaga [25], com concatenação de parâmetros de movimento.	22,29	18,95
H. Cheung e W. Siu [12].	23,53	22,04
Padrão MPEG-4 [13].	21,56	17,28
H. Cheung e W. Siu [13].	21,62	16,75

O método proposto possui resultados razoáveis, se for levado em conta à complexidade dos outros métodos. Por exemplo, o método proposto por A. Smolic [8] que além de necessitar de busca completa, necessita de várias iterações para ser processado. Outro caso é o método proposto por H. Cheung e W. Siu [12] que necessita de busca completa e de gradiente descendente, que é extremamente custoso computacionalmente. Nesses dois casos acima existe o auxílio de máscaras para a segmentação dos objetos para gerar o mosaico. Em relação ao restante dos métodos, existe a necessidade da utilização de gradiente descendente para o refinamento, o que torna os métodos muito complexos. O método proposto é menos complexo, porém deixa a desejar em algumas situações em termos de qualidade.

É possível ver, através da Tabela 1, que a compensação por mosaico tem um desempenho inferior à compensação quadro a quadro, independentemente da técnica utilizada. Também é possível observar que algumas técnicas de compensação por mosaico, mesmo com algoritmos apurados, podem ter uma qualidade de reconstrução inferior, até mesmo, à diferença quadro a quadro. Sendo dessa forma mostrado, que não é uma exclusividade do método proposto o desempenho inferior aos outros dois tipos de compensação. Esta constatação não tira a importância das técnicas de construção do mosaico. Pois, é necessário que as técnicas de construção do mosaico sejam desenvolvidas para a fusão de seqüências naturais com seqüências de vídeos artificiais. Porém, compromete a sua utilização, caso as técnicas atuais de estimação não sejam aprimoradas, para a compressão de vídeo.

Das técnicas de compensação, a clássica busca completa por blocos foi a que obteve melhores resultados. Este fato não invalida as outras técnicas de compensação de movimento, porém ressalta que a busca completa ainda é a forma mais eficaz de compensação de movimento. Outra questão importante é a questão da utilização de técnicas de compensação por mosaico como forma auxiliar de compensação. Pois a estimação por mosaico atua bem em situações em que a compensação pura por blocos não é muito eficaz, como no caso de

grandes translações. Dessa forma, técnicas híbridas de compensação de movimento podem ter um desempenho superior a simples utilização de uma única técnica. Com cada forma de compensação atuando somente nos casos em que seu desempenho for adequado.

Em relação ao método proposto existem alguns pontos falhos como a degradação de qualidade com o aumento do número de quadros. Alguns dos outros métodos apresentam uma degradação muito menor, com o aumento do número de quadros, do que o método proposto. Pois, alguns desses métodos não dependem de valores anteriores de estimação para a estimação atual. Outro ponto negativo do método, comparado com os outros métodos, é a utilização do campo de vetores para a estimação do movimento global. A utilização de um campo de vetores, do modo em que está sendo usado, é falha, pois a técnica de estimação de movimento local não leva em conta o contexto em que o vetor está inserido. Dessa forma o vetor de movimento, muitas vezes, não se comporta como parte do conjunto.

Outra questão importante que contribui para a qualidade do método ser inferior é a falta de segmentação do objeto a frente do fundo. As outras técnicas possuem segmentação dos objetos, que nas duas seqüências de vídeo comparadas, é importante. Portanto, além dos erros devido à estimação, teremos erros gerados pela presença de objetos à frente do fundo, fato que é bastante minimizado nos outros métodos.

## 5. Conclusão

Neste trabalho foi proposto um novo método de estimação de mosaico de uma seqüência de vídeo. Sendo, a estimação de mosaico separada em duas partes, a estimação dos parâmetros de movimento global e a composição do mosaico. A maior contribuição deste método está relacionada à estimação dos parâmetros de movimento global. O método de estimação de movimento global, proposto neste trabalho, possui alguns aspectos diferentes como a filtragem dos vetores de movimento. Mais um ponto que diferencia o método de outros é o tratamento do acúmulo de erros na concatenação dos parâmetros de movimento. Na fusão dos quadros no mosaico foi utilizado um método clássico.

Em termos de resultados para seqüências artificiais, o desempenho foi satisfatório, porém com alguns pontos críticos. O desempenho para as seqüências de vídeo de translação, zoom e uma composição entre as duas transformações o comportamento do método foi muito bom. Mas para seqüências onde ocorre rotação do fundo o desempenho foi muito ruim, o que praticamente inviabiliza o método para este tipo de seqüências de vídeo.

Em comparação com outros métodos, o método proposto neste trabalho obteve um desempenho satisfatório, devido a sua baixa complexidade. Na comparação entre os métodos, para as seqüências de vídeo naturais, é visível que a qualidade do mosaico estimado, utilizando o método proposto, é um pouco inferior. Porém o método não possui uma pré-segmentação de objetos, o que afeta na qualidade final do mosaico construído.

Surpreendentemente a estimação baseada no quadro anterior (com ou sem compensação de movimento) apresentou um desempenho superior, em termos de erro (PSNR), à compensação por mosaico. Isto indica que a compensação baseada no mosaico deve atingir taxa inferior a compressão convencional quadro a quadro. Porém a compensação por mosaico apresenta um desempenho eficaz em algumas situações em que a compensação por blocos não é muito eficiente. Dessa forma, com duas compensações atuando de formas complementares, num

sistema híbrido, poderiam ser mais eficientes do que a utilização de um único tipo de compensação

### **5.1. Trabalhos futuros**

Algumas sugestões para trabalhos futuros são:

- Melhoria, no método proposto, para estimação de movimento global para seqüências vídeo de rotação.
- Aplicação de uma pré-segmentação de objetos no método proposto, para a composição do mosaico.
- Desenvolvimento de uma técnica híbrida de compensação de movimento utilizando procura por blocos e compensação por mosaico.
- Utilização de um outro método de extração dos vetores de movimento para o método, que leve em conta o comportamento global do movimento.
- Comparação entre o tempo de computação do método proposto em relação aos outros métodos de estimação de movimento global.
- Avaliação do desempenho do método proposto, aplicado em um codificador MPEG-4.

## Referências Bibliográficas

- [1] L. Chiariglione, “*Short MPEG-1 description*”, <http://www.chiariglione.org/mpeg/standards/mpeg-1/mpeg-1.htm>, Janeiro 2005.
- [2] L. Chiariglione, “*Short MPEG-2 description*”, <http://www.chiariglione.org/mpeg/standards/mpeg-2/mpeg-2.htm>, Janeiro 2005.
- [3] R. Koenen, “*MPEG-4 Overview – (V.21 – Jeju Version)*”, <http://www.chiariglione.org/mpeg/standards/mpeg-4/mpeg-4.htm>, Janeiro 2005.
- [4] M. Lee, W. Chen, C. B. Lin, C. Gu, T. Markoc, S. I. Zabinsky e R. Szeliski, “*A Layered Video Object Coding System Using Sprite and Affine Motion Model*”. IEEE Trans. Circuits and Syst. Video Technol., vol. 7, pág. 130-145, Fevereiro 1997.
- [5] Y. Lu, W. Gao e F. Wu, “*Efficient Background Video Coding With Static Sprite Generation and Arbitrary-Shape Spatial Prediction Techniques*,” IEEE Trans. Circuits Syst. Video Technol., vol.13, pág. 394-405, Maio 2003.
- [6] M. Irani e P. Anandan, “*Video Indexing Based on Mosaic Representation*”, IEEE Proc., vol. 86, número 5, Maio 1998.
- [7] F. Dufaux, J. Konrad e F. Wu, “*Efficient, robust, and fast global estimation for video coding*,” IEEE Trans. Image Processing, vol.9, pág. 497-501, Mar. 2000.
- [8] A. Smolic, T. Sikora e J.-R. Ohm, “*Long-term global motion estimation and its application for sprite coding, content description and segmentation*”, IEEE Trans. Circuits Syst. Video Technol., vol.9, pág. 1227-1242, Dezembro 1999.
- [9] F. Dufaux e F. Moscheni, “*Motion Estimation Techniques for Digital TV: A Review and a New Contribution*”, IEEE Proc., vol. 83, número 6, Junho 1995.
- [10] C.R. Moloney e E. Dubois, “*Estimation of Motion Field from Image Sequences with Illumination Variation*”, ICASSP-91, vol.4, pág. 2425 – 2428, Abril 1991.
- [11] J. M. Fonseca, “*Acondicionamento de Imagens, operações básicas*”, [www.dee.fct.unl.pt/sdp/disciplinas/ss/20042005/documentos/4%20-%20Acondicionamento.pdf](http://www.dee.fct.unl.pt/sdp/disciplinas/ss/20042005/documentos/4%20-%20Acondicionamento.pdf), Janeiro 2005.
- [12] H. Cheung e W. Siu, “*A Robust Global Motion Estimation Scheme for Sprite Coding*”, ISCAS '03, vol. 2, pág. 632-635, Maio 2003.
- [13] H. Cheung e W. Siu, “*Fast Global Motion Estimation for Sprite Generation*” ISCAS 2002, vol. 3, pág. 1-4, Maio 2003.
- [14] C. R. Wren, “*Perspective Transform Estimation*”, <http://xenia.media.mit.edu/~cwren/interpolator/>, Janeiro 2005.
- [15] F. Zhu, P. Xue e E. Ong, “*Low-Complexity Global Motion Estimation Based on Content Analysis*”, ISCAS '03, vol. 2, pág. 624-627, Maio 2003.
- [16] E. T. Kim e H. Kim, “*Fast and Robust Parameter Estimation Method for Global Motion Estimation in the Video Coder*”, IEEE Trans. Consumer Electronics, vol. 45, pág. 76-83, Fevereiro 1999.

- [17] G. Giunta e U. Mascia, “*Estimation of Global Motion Parameters by Complex Linear Regression*”, IEEE Trans. Image Processing, vol. 8, pág. 1652-1657, Novembro 1999.
- [18] D. Wang e L. Wang, “*Global Motion Parameters Estimation Using a Fast and Robust Algorithm*”, IEEE Trans. Circuits and Syst. Video Technol., vol. 7, pág. 823-826, Outubro 1997.
- [19] T. Chen, W. Lin e C. Chen, “*Artificial Neural Networks for 3-D Motion Analysis –Part I: Rigid Motion*”, IEEE Trans. Neural Networks, vol. 6, pág. 1386-1393, Novembro 1995.
- [20] M.L. Bueno, “*Detecção de Bordas através de Algoritmo Canny*”, <http://www.inf.ufsc.br/~visao/2000/Bordas/>, Fevereiro 2005.
- [21] F. Dufaux e F. Moscheni, “*Background mosaicking for low bit rate coding*”, Proc. IEEE Int. Conf. Image Processing, vol. 1, Lausanne, Switzerland, Setembro 1996, pág. 673-676.
- [22] T. Funkhouser, “*Image Warping*”, <http://www.cs.princeton.edu/courses/archive/fall00/cs426/lectures/warp/>, Janeiro 2005.
- [23] N. Gramaldis, D. Beletsiotis e M.G. Strintzis, “*Sprite Generation and Coding Multiview Image Sequences*”, IEEE Trans. Circuits and Syst. Video Technol., vol. 10, pág. 302-311, Março 2000.
- [24] W. H. Press, B. P. Flannery, S. A. Teukolsky, W. T. Vetterling, “*Numerical Recipes in C: The Art of Scientific Computing*”. Cambridge, U.K.: Cambridge Univ. Press, 1988.
- [25] S. Fukunaga, Y Nakaya, S. H. Son e T. Nagumo, “*MPEG-4 Video Verification Model version 14.2*”, ISO/IEC JTC1/SC29/WG11, pág. 5477, Mauí, Dezembro 1999.