

RECONSTRUÇÃO TRIDIMENSIONAL POR VISÃO ESTÉREO E DETECÇÃO DE OCLUSÕES

Tridimensional Reconstruction through Stereo Vision and Occlusion Detection

Fernando Marques de Almeida Nogueira¹
Clésio Luis Tozzi²

¹**Universidade Federal de Juiz de Fora - UFJF**
Faculdade de Engenharia - FE
Departamento de Engenharia de Produção, Juiz de Fora - MG
Campus Universitário, CEP 36000-000, Juiz de Fora, MG, Brasil.
fernando.nogueira@ufjf.edu.br

²**Universidade Estadual de Campinas - UNICAMP**
Departamento de Engenharia de Computação e Automação Industrial - DCA
Faculdade de Engenharia Elétrica e de Computação – FEEC
CP 6101, CEP 13083-970, Campinas, SP, Brasil.
clesio@dca.fee.unicamp.br

RESUMO

Este trabalho apresenta uma nova abordagem iterativa e hierárquica para a obtenção de mapas densos de disparidade (paralaxe), a partir de um par de imagens estereoscópicas normalizadas, na qual oclusões são explicitamente identificadas. Esta abordagem se fundamenta num processo de busca hierárquica onde são determinados pares homólogos prioritários sobre os quais é aplicada a Injunção de Ordem. Pares homólogos prioritários consistem nos pontos que apresentam a propriedade de Correspondência Bidirecional para o maior número de janelas de busca organizadas em ordem crescente de tamanho. A aplicação da Injunção de Ordem aos pares homólogos prioritários elimina, já nas primeiras iterações, uma grande quantidade de correspondências ambíguas aumentando o índice de acertos e reduzindo o tempo de execução. A abordagem proposta destaca-se pela utilização de apenas dois parâmetros que devem ser fornecidos a priori de maneira empírica: tamanho inicial da janela e número de níveis hierárquicos. A eficiência do algoritmo proposto foi avaliada a partir de imagens reais e simuladas e os resultados comparados com outras abordagens descritas na literatura.

Palavras chaves: Correspondência Estéreo, Oclusão, Disparidade, Paralaxe.

ABSTRACT

This paper presents a new hierarchical and iterative approach to obtain dense disparity (parallax) maps from a pair of normalized stereoscopic images in which occlusions are explicitly identified. This approach is based on a process of hierarchical searching where priority matches are determined and upon which is applied the Order Constraint. Priority matches consist of points that present a property of Bidirectional Correspondence for the greatest number of searching windows, these ones organized in a size ascending order. The Order Constraint application on the priority matches eliminates, usually after the first iterations, a lot of ambiguous matches, increasing the rate of true matches and reducing the execution time. The proposed approach stands out by using only two parameters that must be provided a priori in an empirical way: initial size of window and hierarchical levels number. The efficiency of the proposed algorithm was evaluated using real and simulated images and the results were compared to other approaches described in the literature.

Keywords: Correspondence Stereo, Occlusion, Disparity, Parallax.

1. INTRODUÇÃO

O Problema de Reconstrução consiste em determinar a estrutura tridimensional de uma cena a partir de uma ou mais imagens. Quando duas ou mais imagens da mesma cena tomadas a partir de posições distintas são utilizadas para a obtenção da estrutura tridimensional denomina-se Visão Estéreo.

A Visão Estéreo utiliza a disparidade (paralaxe) entre dois pontos correspondentes situados em imagens distintas, denominados homólogos, como fonte de informação tridimensional da cena. Dessa forma, a principal etapa a ser realizada para solucionar o Problema de Reconstrução por Visão Estéreo consiste em, dado um ponto em uma imagem, determinar o seu homólogo na outra imagem. Esta etapa é conhecida por Problema de Correspondência.

A consideração fundamental utilizada na Visão Estéreo é que a projeção de qualquer região suficientemente pequena da cena é relativamente pouco alterada entre as imagens tomadas em posições distintas. Com isso, a solução do Problema de Correspondência pode ser obtida através do estabelecimento da correspondência por áreas utilizando como primitiva para correlação uma área (janela), resultando em um mapa denso de disparidade.

A utilização de mapas densos de disparidade é motivada por aplicações da Visão Estéreo, tais como inspeção industrial, robótica, cartografia, imagens médicas, entre outras, as quais requerem estimativas dos valores de disparidade para todas as regiões da imagem.

Uma dificuldade inerente ao processo estéreo é a existência nas imagens de pontos de oclusão e para estes os valores de disparidade não podem ser determinados. Pontos de oclusão ocorrem quando um ponto da cena é visível apenas em uma das imagens e são decorrentes das posições distintas das câmeras e das variações de profundidade da cena. Assim, um mapa de disparidade denso deve possuir valores de disparidades para todos os seus pontos, exceto para as regiões do mapa associadas aos pontos de oclusão, que devem ser devidamente rotuladas como tais.

A identificação dos pontos de oclusão se faz importante a fim de minimizar a ocorrência de falsas correspondências entre as imagens estereoscópicas (tais pontos não devem ser considerados no Problema de Correspondência), além de serem importantes fontes de informação para a identificação das regiões do mapa de disparidade onde se observam descontinuidades.

No passado, as pesquisas sobre reconstrução de mapas densos de disparidade tinham por objetivo geral a determinação precisa dos valores de disparidades, embora não inteiramente para todos os pontos do mapa. No entanto, enquanto valores de disparidades imprecisos podem ainda produzir imagens sintetizadas aceitáveis para as regiões de pontos homólogos, regiões de oclusão inconsistentes em geral derivam em erros inaceitáveis.

2. CATEGORIZAÇÃO DE ABORDAGENS EXISTENTES

O Problema de Correspondência pode ser abordado de várias formas. Em geral, as abordagens existentes podem ser classificadas em métodos **locais** ou **globais**.

Os métodos locais utilizam apenas a correspondência baseada em áreas (ou feições) para determinar o valor de disparidade associado a cada *pixel* da imagem de referência. A principal vantagem dos métodos locais é que estes são relativamente simples de serem implementados. Outra vantagem é o baixo tempo de processamento que pode ser obtido utilizando janelas de correlação pequenas ou feições que são facilmente segmentadas. A principal desvantagem desses métodos é a baixa precisão que em geral é observada.

Os métodos globais utilizam em geral uma função-objetivo que deve ser otimizada, cujo domínio é o mapa de disparidade, sendo que todos os valores de disparidades são determinados simultaneamente. A principal vantagem dos métodos globais é a robustez da solução, porém o tempo de processamento é geralmente superior em relação aos métodos locais. Outra grande dificuldade desses métodos está em modelar a própria função-objetivo, uma vez que não existe um modelo matemático já consagrado para este problema. Em geral, os modelos matemáticos utilizados são não-convexos e apresentam muitos mínimos locais, o que torna difícil a obtenção da solução ótima pelos métodos de otimização existentes.

Em termos mais específicos, algumas abordagens classificadas como métodos globais são os métodos baseados em Regularização (Poggio et al., 1985; Horn, 1985; Terzopoulos, 1986), Processos Cooperativos (Zitnick & Kanade, 2000), Programação Dinâmica (Ohta & Kanade, 1985 e Belhumeur, 1993), Fluxo Máximo em Grafos (Roy & Cox, 1998; Ishikawa & Geiger, 2001; Boykov et al., 2001 e Kolmogorov & Zabih, 2001), *Layered* (Lin & Tomasi, 2002; Darrell and A. Pentland, 1995) e Segmentos (Bleyer & Gelautz, 2005; Deng et al, 2007).

De maneira cronológica, os métodos locais foram inicialmente propostos seguidos dos métodos globais. Naturalmente, os métodos mais recentes têm apresentado melhores resultados em comparação aos métodos propostos anteriormente.

No entanto, percebe-se que paralelamente a esta evolução, é cada vez maior o número de parâmetros utilizados nas abordagens propostas. Em geral, tais parâmetros são determinados de maneira empírica além de serem cada vez mais dependentes das características das imagens e/ou das cenas. Com isso, um conjunto de valores designados para os parâmetros e que resultam em um mapa de disparidade com ótima qualidade para um par de imagens pode resultar em um mapa de disparidade de baixa qualidade para outro par de imagens

estereoscópicas, não obstante a utilização da mesma abordagem.

O fato citado acima coloca em questão a utilização das abordagens que possuem uma grande quantidade de parâmetros determinados empiricamente e extremamente sensíveis às condições e características das imagens.

Entende-se que uma boa abordagem é aquela que obtém bons resultados para um grande número de imagens estereoscópicas com características bastante distintas entre elas e que utilizam valores relativamente constantes para os parâmetros empregados na abordagem, independente do número de parâmetros que possua.

Ainda, em relação à identificação de oclusões, esta é diretamente dependente de um limiar (parâmetro) nos métodos citados acima, o qual em geral, é determinado de maneira também empírica.

3. ABORDAGEM PROPOSTA

Considerando as dificuldades dos métodos globais, este trabalho propõe uma abordagem baseada em um processo hierárquico e número reduzido de parâmetros de controle e que, mesmo considerando um tratamento local, obtém resultados equiparáveis aos obtidos por complexos sistemas globais, possibilitando adicionalmente a determinação de oclusões.

A abordagem proposta opera de maneira local entre os pares de linhas epipolares homólogas, uma vez que os pares de linhas epipolares homólogas são tratados de maneira independente entre si, e opera de maneira global entre os *pixels* sobre um mesmo par de linhas epipolares homólogas. Dessa forma, esta abordagem não pode ser classificada formalmente como um método local ou global. Entende-se que o correto é classificar a abordagem proposta em uma nova categoria, proposta neste trabalho, como método híbrido.

A abordagem proposta neste trabalho pode ser considerada também um método baseado em Processos Cooperativos, uma vez que a Injunção de Ordem utilizada define uma região inibitória e uma região excitatória, porém de maneira booleana, excluindo ou não pares de *pixels* candidatos a pares homólogos.

3.1 Matriz de Correspondência

Seja um par de imagens estereoscópicas de dimensões $H \times W$, onde H é o número de linhas e W é o número de colunas, uma matriz M com dimensões $H \times W \times H \times W$ pode ser montada, na qual o valor do elemento $M(x_E, y_E, x_D, y_D)$ representa uma medida de correspondência entre um ponto $p_E(x_E, y_E)$ na imagem esquerda e um ponto $p_D(x_D, y_D)$ na imagem direita.

Sem perda de generalidade, o par de imagens estereoscópicas pode ser retificado através da reamostragem das linhas das imagens ao longo das linhas epipolares. Este processo é denominado Retificação Estéreo ou Normalização (Nogueira & Tozzi (1998)), o

qual permite reduzir uma dimensão da Matriz de Correspondência.

Desta forma para cada linha epipolar i tem-se uma matriz $W \times W$ e, portanto, para todas as linhas epipolares tem-se H matrizes $W \times W$. O valor associado a cada posição da Matriz de Correspondência $M(x_E, x_D, y)$ com dimensões $H \times W \times W$, representa uma medida de correspondência entre um ponto $p_E(x_E, y)$ na imagem esquerda retificada e um ponto $p_D(x_D, y)$ na imagem direita retificada.

A figura 1 mostra duas câmeras com $W = 7$ para uma linha epipolar i , a figura 2 mostra sua respectiva Matriz de Correspondência à esquerda e à direita a Matriz de Correspondência para $H=3$.

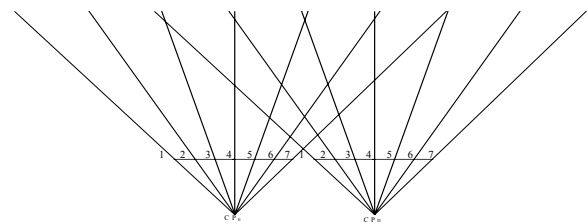


Fig. 1 – Câmeras com $W = 7$.

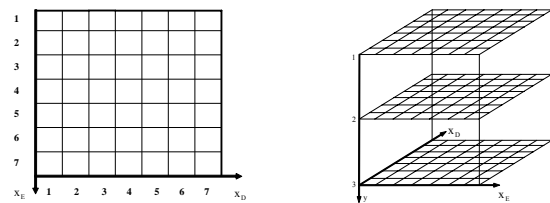


Fig. 2 – Matriz de Correspondência para uma linha epipolar (esq.) e para $H=3$ (dir.).

3.2 Injunções e seus Efeitos na Matriz de Correspondência

Todas as injunções apresentadas nesta seção são impostas independentemente para cada par de linhas epipolares homólogas i . Assim, o efeito de cada injunção sobre a Matriz de Correspondência pode ser visualizado em matrizes com dimensões $W \times W$.

Injunções entre linhas epipolares não homólogas não são tratadas neste trabalho.

3.2.1 Disparidade Não Negativa

O processo de Normalização torna as imagens estereoscópicas análogas às imagens capturadas com a seguinte configuração de câmeras:

- 1) distância focal da câmera esquerda igual à distância focal da câmera da direita; e
- 2) planos imagem da câmera esquerda e da direita paralelo ao eixo epipolar (base).

Com as condições 1 e 2 citadas acima, a disparidade d entre um par de pontos homólogos $p_E(x_E, y)$ e $p_D(x_D, y)$ é dada por:

$$d = x_E - x_D \quad (1)$$

A profundidade Z do ponto associado no espaço objeto é dada por:

$$Z = \frac{b \cdot f}{d} \quad (2)$$

onde:

b é o comprimento da base; e

f é a distância focal das câmeras da esquerda e da direita.

Para profundidades no intervalo $[0, +\infty]$ e considerando que os parâmetros b e f na expressão 2 são sempre positivos, os valores de disparidade d também estão no intervalo $[0, +\infty]$ (não negativo). Valores de disparidades negativas resultariam em profundidades negativas, portanto atrás das câmeras, o que é impossível.

De acordo com a expressão 1, pode-se concluir então que:

$$x_E \geq x_D \quad (3)$$

A figura 3 mostra o efeito dessa injunção na Matriz de Correspondência M.

1		X	X	X	X	X	X
2			X	X	X	X	X
3				X	X	X	X
4					X	X	X
5						X	X
6							X
7							

X = pares impossíveis

Fig. 3 – Efeito da Injunção de Disparidade Não Negativa.

A utilização desta injunção reduz significativamente o custo computacional. Considerando que as restrições são cumulativas, as próximas figuras que demonstram os efeitos das demais injunções estarão acrescidas desta.

3.2.2 Ordem

A injunção de ordem impõe que a ordem dos pontos homólogos não seja alterada. Seja $p_E(k_E, y)$ e $p_D(k_D, y)$ um par de pontos homólogos e $L \in \mathfrak{R}^+$, então:

$$p_E(k_E + L, y) \text{ não pode ser homólogo a qualquer } p_D(q, y), \text{ com } q \leq k_D \quad (4)$$

$$p_E(k_E - L, y) \text{ não pode ser homólogo a qualquer } p_D(q, y), \text{ com } q \geq k_D \quad (5)$$

$$p_D(k_D + L, y) \text{ não pode ser homólogo a qualquer } p_E(q, y), \text{ com } q \leq k_E \quad (6)$$

$$p_D(k_D - L, y) \text{ não pode ser homólogo a qualquer } p_E(q, y), \text{ com } q \geq k_E \quad (7)$$

A figura 4 mostra o efeito dessa injunção na representação esquemática das câmeras (como exemplo, $(k_E, k_D) = (6, 2)$). O ponto preto representa o par de pontos homólogos escolhido, e os pontos cinza representam os pares impossíveis decorrentes desta escolha e da Injunção de Ordem.

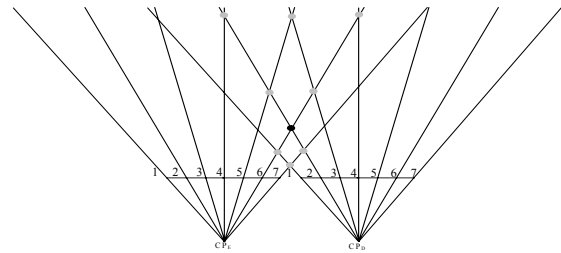


Fig. 4 – Injunção de Ordem.

A figura 5 mostra o efeito conjunto dessa injunção na Matriz de Correspondência M e da injunção de Disparidade Não Negativa.

1		X	X	X	X	X	X
2		W	X	X	X	X	X
3		W	W	X	X	X	X
4		W	W	W	X	X	X
5		W	W	W	W	X	X
6		W	O	W	W	W	X
7		W	W				

X = Pares impossíveis pela Injunção de Disparidade Não-Negativa
O = Par Escolhido

W = Pares impossíveis pela Injunção de Ordem

Fig. 5 – Efeito conjunto da Injunção de Ordem e de Disparidade Não Negativa.

Esta injunção não é válida para regiões da cena com superfícies transparentes e/ou regiões onde o módulo do gradiente de disparidade é maior ou igual a 2.

3.2.3 Limite do Gradiente de Disparidade

O Módulo do Gradiente de Disparidade $|\gamma|$ pode ser definido como a diferença entre as disparidades associadas a dois pares de pontos homólogos próximos (1 e 2), divididos por sua separação (distância) no sistema de coordenadas do espaço ciclopeano. Analiticamente:

$$|\gamma| = \frac{|d_2 - d_1|}{|D_{12}|} \quad (8)$$

onde:

d_1 e d_2 são valores de disparidade associados aos pontos 1 e 2;

D_{12} é a distância entre 1 e 2 no espaço ciclopeano.

Considerando imagens retificadas, a transformação do sistema de coordenadas do espaço das imagens estereoscópicas para o sistema de coordenadas do espaço ciclopeano é dada por:

$$x_C = \frac{x_E + x_D}{2} \quad (9)$$

$$y_C = y_E = y_D \quad (10)$$

$$d(x_C, y_C) = x_E - x_D \quad (11)$$

A transformação do sistema de coordenadas do espaço ciclopeano para o sistema de coordenadas do espaço das imagens estereoscópicas é dada por:

$$x_E = x_C + \frac{d(x_C, y_C)}{2} \quad (12)$$

$$y_C = y_E = y_D \quad (13)$$

$$d(x_C, y_C) = x_E - x_D \quad (14)$$

Pode-se verificar a equivalência entre a Injunção de Ordem e o Limite do Módulo do Gradiente de Disparidade igual a 2 através de: seja $x_E = k_{E1}$ a coordenada na direção x de um ponto 1 na imagem esquerda e $x_D = k_{D1}$ a coordenada na direção x deste mesmo ponto 1, porém na imagem direita e $x_E = k_{E2}$ e $x_D = k_{D2}$ de maneira análoga ao ponto 1, porém para o ponto 2. Substituindo estas coordenadas na expressão 8, fica:

$$|\gamma| = \frac{|d_2 - d_1|}{|D_{12}|} = \frac{|(k_{E2} - k_{D2}) - (k_{E1} - k_{D1})|}{\left| \frac{(k_{E1} + k_{D1}) - (k_{E2} + k_{D2})}{2} \right|} \quad (15)$$

Fazendo simples manipulações algébricas tem-se que:

$$|\gamma| = \frac{|k_{E2} - k_{D2} - k_{E1} + k_{D1}|}{\left| \frac{(k_{E1} + k_{D1}) - (k_{E2} + k_{D2})}{2} \right|} = 2 \frac{|k_{E2} - k_{D2} - k_{E1} + k_{D1}|}{|k_{E1} + k_{D1} - k_{E2} - k_{D2}|} \quad (16)$$

De acordo com Pollard et. al., 1986, a expressão 16 pode-se ser rearranjada para melhor compreensão:

$$|\gamma| = 2 \frac{|(k_{D1} - k_{D2}) - (k_{E1} - k_{E2})|}{|(k_{D1} - k_{D2}) + (k_{E1} - k_{E2})|} \quad (17)$$

As coordenadas do ponto 2 podem ser definidas em função das coordenadas do ponto 1 por:

$$k_{E2} = k_{E1} + M \quad (18)$$

$$k_{D2} = k_{D1} + N$$

onde:

$$M \in \mathfrak{R}^+$$

$$N \in \mathfrak{R}^+$$

Substituindo 18 em 17, fica:

$$|\gamma| = 2 \frac{|(k_{D1} - (k_{D1} + N)) - (k_{E1} - (k_{E1} + M))|}{|(k_{D1} - (k_{D1} + N)) + (k_{E1} - (k_{E1} + M))|} = 2 \frac{|M - N|}{|-M - N|} \quad (19)$$

A inversão de ordem ocorre quando:

$$M > 0 \quad \text{e} \quad N < 0 \quad (20)$$

ou

$$M < 0 \quad \text{e} \quad N > 0 \quad (21)$$

Analisando os intervalos dados em 20 e 21 em 19, verifica-se que para estes o quociente $\frac{|M - N|}{|-M - N|} > 1$, o que implica que quando $|\gamma| > 2$ ocorre a inversão de ordem.

4. DETECÇÃO DE OCLUSÃO - CORRESPONDÊNCIA BIDIRECIONAL E INJUNÇÃO DE ORDEM

O problema de detecção de oclusão é resolvido através da utilização conjunta da Correspondência Bidirecional e da imposição da Injunção de Ordem.

A Correspondência Bidirecional consiste em analisar as medidas de correspondência da imagem da esquerda em relação à imagem da direita e vice-versa.

O princípio que sustenta a utilização da Correspondência Bidirecional para a detecção de oclusão é que um par de *pixels* homólogos p e p' deve apresentar Correspondência Bidirecional, ou seja, a melhor medida de correspondência para um *pixel* p na imagem da esquerda é com um *pixel* p' na imagem da direita, que por sua vez, possui melhor medida de correspondência para p.

Desconsiderando a existência de superfícies transparentes na cena e uma vez que oclusões são *pixels* na imagem da esquerda ou da direita que não possuem homólogos, as linhas ou colunas na Matriz de Correspondência que não possuem pares correspondentes bidirecionais são "fortes candidatos" a serem classificados como oclusões.

Utilizando a Matriz de Correspondência, a Correspondência Bidirecional é implementada realizando

a busca das melhores medidas de Correspondência ao longo das linhas e colunas da Matriz de Correspondência.

Apesar da Correspondência Bidirecional possibilitar a identificação de oclusões, esta heurística é utilizada neste trabalho com o intuito de determinar com melhor precisão os pares homólogos, e não os *pixels* que não possuem homólogos (oclusões).

A utilização da Injunção de Ordem, ou de maneira análoga, a utilização de um limite para o módulo do Gradiente de Disparidade igual a 2, como já citado no item 3.2.3, impõe que oclusões ocorram sempre que houver variação dos valores de disparidade associados a *pixels* vizinhos.

Utilizando a notação empregada na seção 3.2.3, onde $x_E = k_{E1}$ e $x_D = k_{D1}$ são as coordenadas na direção x do par 1 na imagem da esquerda e da direita, respectivamente e $x_E = k_{E2}$ e $x_D = k_{D2}$ de maneira análoga, porém para o par 2, as coordenadas deste par 2 podem ser definidas em função das coordenadas do par 1 por:

$$\begin{aligned} k_{E2} &= k_{E1} + M \\ k_{D2} &= k_{D1} + N \end{aligned} \quad (22)$$

onde:

$$\begin{aligned} M &\in \mathfrak{R}^+ \\ N &\in \mathfrak{R}^+ \end{aligned}$$

Assim, para um par de *pixels* vizinhos 1 e 2, com disparidades d_1 e d_2 , respectivamente, a diferença $\Delta d = d_1 - d_2$ fica:

$$\Delta d = (k_{E1} - k_{D1}) - (k_{E2} - k_{D2}) \quad (23)$$

Aplicando 22 em 23 tem-se que:

$$\Delta d = (k_{E1} - k_{D1}) - (k_{E1} + M - (k_{D1} + N)) \quad (24)$$

que implica em:

$$\Delta d = -M + N \quad (25)$$

Para imagens discretas e considerando dois *pixels* vizinhos na imagem da esquerda, M assume o valor 1, assim:

$$\Delta d = -1 + N \quad (26)$$

A expressão 26 mostra que respeitando a Injunção de Ordem, faz-se necessário existir uma região com $N-1$ *pixels* de oclusão na imagem da direita se a diferença das disparidades associadas a estes *pixels* vizinhos é Δd . Com isso, pode-se deduzir um importante resultado: "a diferença de disparidade $\Delta d = 1$ na imagem da esquerda entre dois *pixels* vizinhos, implica em $N = 2$ para satisfazer a igualdade em 26 e, portanto, existe $N-1 = 2-1 = 1$ *pixel* na imagem da direita que é oclusão". Para

outros valores de Δd e para oclusões na imagem da esquerda o raciocínio é análogo.

Em termos práticos, o resultado deduzido acima mostra que sempre que houver variação dos valores de disparidade associados a dois *pixels* vizinhos em uma imagem ocorre uma oclusão na outra imagem cujo tamanho da região de oclusão é igual a esta variação dos valores de disparidade.

5. TAMANHO DA JANELA DE CORRELAÇÃO - CORRESPONDÊNCIA BIDIRECIONAL HIERÁRQUICA

Existem várias possibilidades para se determinar um critério que defina sobre quais pares deve-se impor a Injunção de Ordem. Um critério simples é escolher qual a melhor medida de correspondência associada a cada *pixel*, por exemplo, na imagem da esquerda e então impor a Injunção de Ordem para estes *pixels*. Utilizando a Matriz de Correspondência, este critério se dá analisando qual a melhor medida de correspondência ao longo das linhas dessa matriz.

Embora simples este critério não é robusto, uma vez que os resultados obtidos para os *pixels* na imagem da esquerda (analisando a Matriz de Correspondência ao longo das linhas dessa matriz) podem diferir dos resultados obtidos para os *pixels* na imagem da direita (analisando a Matriz de Correspondência ao longo das colunas dessa matriz).

Um critério mais robusto é escolher como *pixels* homólogos aqueles que apresentam a propriedade de Correspondência Bidirecional (como citado na seção anterior) e então impor a Injunção de Ordem para estes pares de *pixels*.

Utilizando a técnica de Correspondência baseada em Áreas, as medidas de Correspondência são, entre outras, função do tamanho da janela de correspondência utilizada. Assim, o problema do tamanho da janela torna-se evidente.

Janelas pequenas são necessárias em regiões onde a superfície da cena não é fronto-planar, entretanto, a pequena quantidade de *pixels* (amostras) contidas na janela tende a aumentar a ambigüidade das medidas de correspondência. Janelas grandes possuem uma maior quantidade de *pixels* (amostras), o que tende a minimizar as ambigüidades devido às medidas de correlação serem mais precisas, porém, a posição de melhor medida de correspondência pode não resultar em um valor de disparidade correto devido à janela cobrir uma região referente a superfícies não necessariamente fronto-planar.

A maneira proposta neste trabalho de minimizar o problema do tamanho da janela é quantificar os pares que possuem a propriedade de Correspondência Bidirecional a partir de medidas de correspondências oriundas de vários tamanhos de janela.

A heurística desta abordagem é que um par que possui a propriedade de Correspondência Bidirecional a partir de medidas oriundas de vários tamanhos de janela

deve ser provavelmente formado por *pixels* homólogos de que aqueles que não são.

De maneira resumida, a heurística citada acima diz que dois *pixels* são homólogos se estes possuem a propriedade de Correspondência Bidirecional independente do tamanho da janela de correlação. No entanto, esta heurística não considera o problema de preservar as descontinuidades da cena.

Para tentar preservar as descontinuidades da cena, pode-se quantificar quais os pares que são Correspondências Bidirecionais primeiramente para as medidas oriundas do menor tamanho de janela utilizado e então, posteriormente, para os demais tamanhos de janelas organizadas de maneira hierárquica em ordem crescente, de tal forma que um par apenas poderá ser Correspondência Bidirecional se este mesmo par é também Correspondência Bidirecional para as medidas oriundas das janelas menores. Esta organização hierárquica impõe o conceito de níveis de informações, onde os níveis inferiores correspondem às medidas oriundas das janelas menores e os níveis superiores às medidas oriundas das janelas maiores.

Esta hierarquia é do tipo "*bottom - up*" e esta foi adotada (e não ao contrário, do tipo "*top-down*") porque nas regiões de descontinuidades as janelas pequenas sofrem muito menos o efeito da variação local da disparidade. A desvantagem em utilizar janelas pequenas é o aumento do número de ambigüidades dificultando a resolução do Problema de Correspondência, como já citado. A fim de minimizar esta dificuldade, ou seja, reduzir o número de ambigüidades preservando as descontinuidades, é que se utilizam janelas maiores de maneira hierárquica.

É de se esperar que os pares que possuem a propriedade de Correspondência Bidirecional em um número maior de níveis hierárquicos tenham maiores chances de estarem correto (ser realmente formado por *pixels* homólogos) de que aqueles que são Correspondência Bidirecional em apenas poucos níveis hierárquicos. Com isso, deve-se então aplicar a Injunção de Ordem sobre os pares com as maiores chances entre todos os demais.

Ao impor a Injunção de Ordem para os pares com maiores prioridades, restarão provavelmente ainda alguns pares com prioridades menores e que não foram classificados como pares impossíveis em decorrência da imposição da Injunção de Ordem para os pares com maiores prioridades. No entanto, a imposição da Injunção de Ordem para os pares com maiores prioridades pode fazer com que a prioridade de outros pares com prioridades mais baixas aumente em uma nova verificação das prioridades. Isto ocorre porque a imposição da Injunção de Ordem sobre um par classifica outros pares como impossíveis.

Após a imposição da Injunção de Ordem sobre os pares com maiores prioridades faz-se necessário verificar se ainda existem pares não resolvidos. Caso ainda existam pares não resolvidos, faz-se necessário reavaliar a Correspondência Bidirecional para estes pares

e quantificar as suas prioridades para então aplicar a Injunção de Ordem novamente. Este fato impõe a necessidade de um processo iterativo.

O fato citado pode ser visto também como uma restrição do espaço de busca do *pixel* homólogo devido à ocorrência de outro par, o que torna a abordagem proposta um Processo Cooperativo, uma vez que a ocorrência de um par de *pixels* homólogos favorece a ocorrência de outros pares na região que a Injunção de Ordem é respeitada enquanto elimina a ocorrência dos pares na região que esta injunção é violada.

Com isso, o processo iterativo termina quando não houver mais pares não resolvidos (ainda ambíguos) para um par de linhas epipolares homólogas.

6. MATRIZ DE PRIORIDADES E PROCESSO ITERATIVO

A fim de organizar as prioridades de cada par, uma Matriz de Prioridades P, com a mesma estrutura de uma Matriz de Correspondências, porém com uma dimensão a menos, é montada, onde cada elemento desta matriz possui uma medida de prioridade para se impor a Injunção de Ordem.

A hierarquia do tipo "*bottom - up*" é descrita da seguinte maneira: "um elemento $P(x_E, x_D)$ na Matriz de Prioridades, cuja medida de correspondência entre x_E e x_D para um nível hierárquico $n + 1$ resulte em uma Correspondência Bidirecional, somente terá prioridade $n + 1$ se a medida de correspondência para estes mesmos *pixels* para o nível hierárquico n também resulte em uma Correspondência Bidirecional".

O processo hierárquico descrito necessita ser realizado de maneira iterativa, no qual, a cada iteração são selecionados os pares de *pixels* com maiores prioridades e então imposta a Injunção de Ordem sobre estes até que não existam mais pares ambíguos em um par de linhas epipolares homólogas.

É importante ressaltar também que a correspondência baseada em Áreas, geralmente um processo de alto custo computacional, é realizada apenas uma vez, embora para vários tamanhos de janela. A cada iteração as Matrizes de Correspondência são apenas atualizadas com rótulos que representam pares impossíveis devido à imposição da Injunção de Ordem.

A principal idéia que suporta este processo iterativo é que para cada par verdadeiro existem vários outros pares falsos, de tal forma que a ocorrência de um par verdadeiro auxilia de maneira cooperativa os demais pares verdadeiros ainda ambíguos, enquanto que ao mesmo tempo desfavorece os pares falsos, porém ainda ambíguos e que poderiam vir a ser classificados erroneamente como pares verdadeiros.

Apenas com o intuito de formalização, pode-se definir a seguinte regra para o evento descrito:

Regra 1

Selecionar pares correspondentes e impor a Injunção de Ordem para os pares com maiores prioridades em uma iteração.

7. REGRAS COMPLEMENTARES

A regra 1, embora básica, é insuficiente para estabelecer completamente a abordagem proposta. Em vista deste fato, algumas regras complementares devem ser estabelecidas.

O principal problema que ocorre é quando um par com máxima prioridade em uma iteração está localizado na região de pares impossíveis determinada por outro par também com máxima prioridade na mesma iteração. Neste caso a imposição da Injunção de Ordem sobre um par implica na classificação do outro como par impossível. Este fato ocorre porque as prioridades são grandezas discretas, o que possibilita haver empate de dois ou mais pares localizados em regiões conflitantes com a mesma prioridade máxima em uma iteração.

A fim de sanar este problema, a seguinte regra pode ser utilizada:

Regra 2

No caso em que dois ou mais pares com prioridade máxima em uma iteração ocorram em regiões conflitantes deve-se diminuir a prioridade de -1 daquele que para ocorrer seja necessário classificar como par impossível o maior número de pares com prioridade máxima.

Em caso de empate do número de pares que ocorrem na região conflitante, a seguinte regra é utilizada:

Regra 3

Diminuir a prioridade de -1 daquele par que tiver maior disparidade (menor profundidade).

A regra 3 não é suportada por nenhuma idéia de caráter lógico e deve ser entendida apenas como uma decisão que tem que ser tomada quando os demais critérios de decisão empataram.

8. ALGORITMO E IMPLEMENTAÇÃO

A medida de correspondência empregada nesta abordagem é a Somatória das Diferenças Absolutas (SDA) entre uma janela na imagem esquerda e uma janela na imagem direita. Esta medida de correspondência, a qual formalmente é uma medida de correlação, foi adotada por ser bastante simples e apresentar um menor custo computacional em relação a várias outras medidas de correlação, tais como, Somatória das Diferenças Quadráticas ou Coeficiente de Correlação. A SDA é dada por:

$$SDA(y, x_E, x_D) = \sum_{i=-\frac{S_i}{2}}^{\frac{S_i}{2}} \sum_{j=-\frac{S_j}{2}}^{\frac{S_j}{2}} |p_E(y+i, x_E+j) - p_D(y+i, x_D+j)| \quad (27)$$

onde:

S_i é o tamanho da janela na direção i ;

S_j é o tamanho da janela na direção j ;

$p_E(\cdot, \cdot)$ e $p_D(\cdot, \cdot)$ são os valores de intensidade dos *pixels* na imagem da esquerda e da direita, respectivamente.

Utilizando imagens com formato 8 *bits* e $S_i = S_j = 1$, o maior valor para SDA é 255 e a menor é 0 (zero). Cabe ressaltar que a medida 0 (zero) implica em melhor medida possível e 255 em pior medida possível.

Tais medidas $SDA(y, x_E, x_D)$ podem ser normalizadas para o intervalo $[0, 1]$ fazendo:

$$SDA^n(y, x_E, x_D) = \frac{SDA(y, x_E, x_D)}{255 \times S_i \times S_j} \quad (28)$$

onde:

$SDA^n(i, j_E, j_D)$ é a medida $SDA(i, j_E, j_D)$ normalizada em $[0, 1]$.

Dado um tamanho de janela inicial $S_i^1 \times S_j^1$, e considerando níveis hierárquicos, a correlação deve ser calculada utilizando janelas com tamanhos $S_i^1 \times S_j^1$, $S_i^1 + 2 \times S_j^1 + 2$, $S_i^1 + 4 \times S_j^1 + 4, \dots, S_i^1 + 2(N-1) \times S_j^1 + 2(N-1)$, onde N é o nível máximo do processo hierárquico. A SDA fica:

$$SDA^n(n, y, x_E, x_D) = \sum_{i=-\frac{S_i^n}{2}}^{\frac{S_i^n}{2}} \sum_{j=-\frac{S_j^n}{2}}^{\frac{S_j^n}{2}} |p_E(y+i, x_E+j) - p_D(y+i, x_D+j)| \quad (29)$$

onde:

$$S_i^n = S_i^1 + 2(n-1) \quad (30)$$

$$S_j^n = S_j^1 + 2(n-1) \quad (31)$$

$$n \in \mathbb{Z} \mid 1 \leq n \leq N \quad (32)$$

A seguinte simbologia é utilizada no fluxograma abaixo:

n → número do nível hierárquico;

N → número de níveis hierárquicos;

$$S_i^n \rightarrow S_i^1 + 2(n-1)$$

$$S_j^n \rightarrow S_j^1 + 2(n-1)$$

h → altura das imagens estereoscópicas;

w → largura das imagens estereoscópicas;

y → número do par de linhas epipolares;

$$b \rightarrow \frac{S_i^N - 1}{2};$$

$x_E \rightarrow$ número da coluna na imagem da esquerda de um *pixel* no intervalo $[b-1, w-b-1]$;

$x_D \rightarrow$ número da coluna na imagem da direita de um *pixel* no intervalo $[b-1, w-b-1]$;

$x_E^\Psi \rightarrow$ número da coluna na imagem da esquerda de um *pixel* contido na região de pares impossíveis determinada por (x_E, x_D) ;

$x_D^\Psi \rightarrow$ número da coluna na imagem da direita de um *pixel* contido na região de pares impossíveis determinada por (x_E, x_D) ;

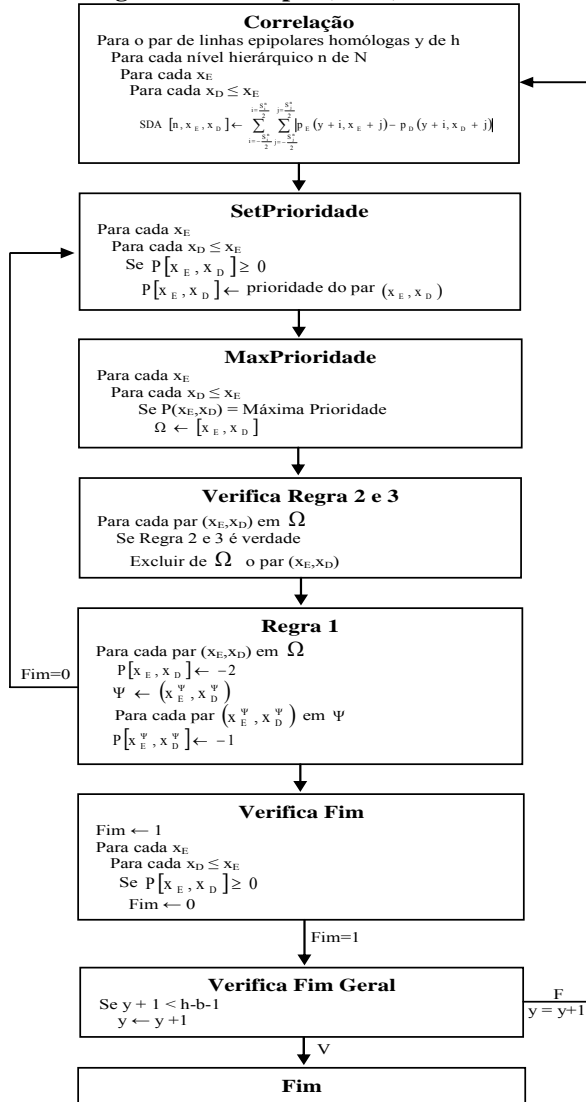
$p_E(\cdot, \cdot)$ e $p_D(\cdot, \cdot) \rightarrow$ valores de intensidade dos *pixels* na imagem da esquerda e da direita, respectivamente.

Ω e $\Psi \rightarrow$ conjunto de pares (x_E, x_D) ;

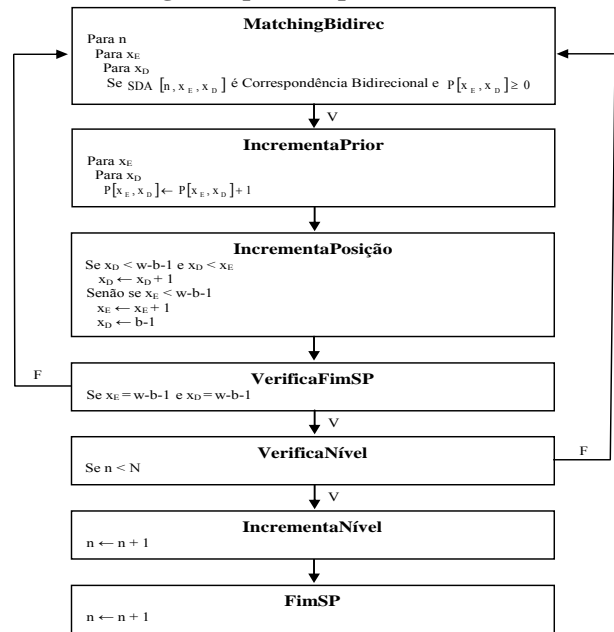
$SDA[\cdot, \cdot, \cdot] \rightarrow$ matriz com medidas de correlação;

$P[\cdot, \cdot] \rightarrow$ matriz com valores de prioridades.

Algoritmo Principal (Main):



Fluxograma para Setprioridade:



9. TESTES E RESULTADOS

As imagens utilizadas nos testes realizados são normalizadas e não contém objetos transparentes. No entanto, muitas delas apresentam regiões com texturas homogêneas e variações abruptas (não suave) dos valores de disparidade, o que ocasionam grandes regiões de oclusão, dificultando substancialmente a resolução do Problema de Correspondência.

Com o intuito de quantificar os resultados obtidos, utilizou-se um conjunto de imagens estereoscópicas disponíveis no site <http://www.middlebury.edu/stereo>. Este conjunto contém, além de imagens estereoscópicas reais (Map e Tsukuba) e sintéticas (Sawtooth e Venus), os respectivos mapas de disparidade verdadeiros, mapas de regiões de oclusão e mapas de regiões sem texturas, sendo todos estes associados às imagens da esquerda.

Com o intuito de facilitar a comparação dos resultados com outras abordagens utilizou-se a mesma métrica elaborada por Scharstein & Szeliski, 2002, na qual um par é considerado errado se o módulo da diferença entre o valor de disparidade calculado e o valor de disparidade obtido do mapa de disparidade verdadeiro é maior que 1.

O resultado total é contabilizado como a porcentagem de pares errados, em relação ao total de *pixels* nas imagens estereoscópicas descontando uma "borda" nas regiões limítrofes das imagens estereoscópicas e os *pixels* contidos nas regiões de oclusões oriundos dos mapas de regiões de oclusão verdadeiros.

As figuras 6, 8, 10 e 12 mostram, à esquerda, as imagens estereoscópicas da esquerda normalizadas e à direita, os mapas de pares errôneos.

As figuras 7, 9, 11 e 13 mostram os mapas de disparidade reconstruídos (imagens à esquerda) e os mapas de disparidades verdadeiras (imagens à direita). Nestas imagens as regiões de oclusão foram rotuladas com valores 255 e estão representadas em branco. As regiões com os demais níveis de cinza codificam os valores de disparidade, onde as regiões claras representam altos valores de disparidade e as regiões escuras representam baixos valores de disparidade.

Todos os testes foram realizados com 5 níveis hierárquicos, porém para 7 valores distintos do tamanho inicial da janela de correlação (exemplificando, se $S_i^1 \times S_j^1 = 3 \times 3$, tem-se $S_i^5 \times S_j^5 = 11 \times 11$).

As tabelas 1 (Map), 2 (Sawtooth), 3 (Tsukuba) e 4 (Venus) apresentam resultados quantitativos das reconstruções em função dos valores do tamanho inicial da janela de correlação, onde E é a porcentagem de pares, em relação ao total de pares, cujos valores dos mapas de disparidades reconstruídos diferem, em módulo, de valores maiores que 1 dos respectivos valores dos mapas de disparidades verdadeiros; ET é análogo à E, porém calculado apenas para as regiões classificadas como “sem texturas”, EO é análogo à E, porém calculado apenas para as regiões de oclusão e T é o tempo de processamento em segundos utilizado apenas como referência para estimativa de esforço computacional decorrente do aumento do tamanho inicial da janela e determinados para o programa desenvolvido e executado em um processador Pentium 4, 2.8GHz e 512 Mb RAM. Ainda para estas tabelas, os valores em negrito representam os melhores resultados.

Os mapas de pares errados (figuras 6(dir.), 8(dir.), 10(dir.) e 12(dir.)) ilustram a ocorrência espacial destes pares.

Os mapas de disparidades reconstruídos apresentados são referentes aos melhores resultados obtidos para cada par de imagens utilizadas entre os vários valores de tamanho da janela inicial de correspondência.

9.1 Imagens Map

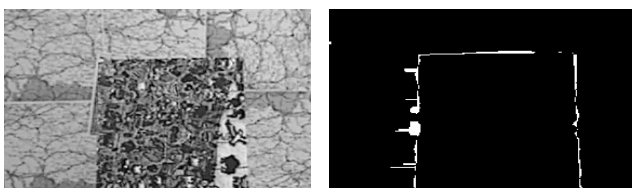


Fig. 6 – Imagem esquerda (esq.) e Mapa de pares errados para Map (dir.).

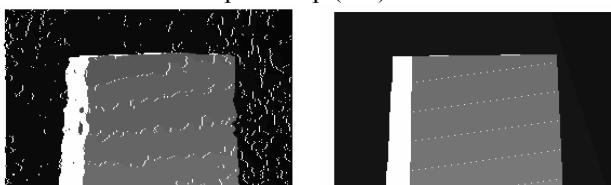


Fig. 7 – Mapas de disparidades esquerdos com oclusões reconstruído (esq.) e verdadeiro (dir.) para Map.

TABELA 1 – RESULTADOS PARA “MAP”.

	Tamanho da janela inicial						
	1x1	3x3	5x5	7x7	9x9	11x11	13x13
E(%)	18.6	3.4	0.9	0.7	0.8	0.9	1.2
ET(%)	34.8	32.9	5.9	0.7	0.2	0.5	0.7
EO(%)	15.7	9.9	8.5	9.6	10.7	12.2	13.3
T(s)	22	6	7	8	10	12	14

9.2 Imagens Sawtooth

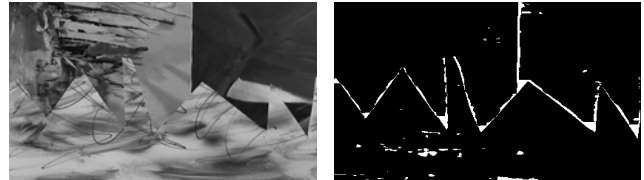


Fig. 8 – Imagem esquerda (esq.) e Mapa de pares errados para Sawtooth (dir.).

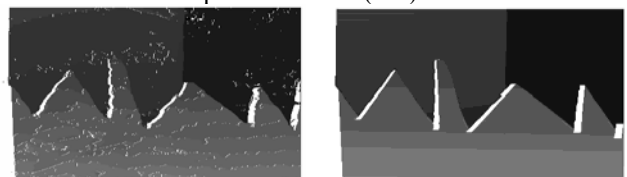


Fig. 9 – Mapas de disparidades esquerdos com oclusões reconstruído (esq.) e verdadeiro (dir.) para Sawtooth.

TABELA 2 – RESULTADOS PARA “SAWTOOTH”.

	Tamanho da janela inicial						
	1x1	3x3	5x5	7x7	9x9	11x11	13x13
E(%)	20.7	5.1	2.6	2.2	2.3	2.7	3.0
ET(%)	26.2	11.6	2.7	0.9	0.8	1.5	2.3
EO(%)	20.8	15.3	16.6	19.9	23.1	25.8	27.9
T(s)	177	17	16	18	21	24	28

9.3 Imagens Tsukuba

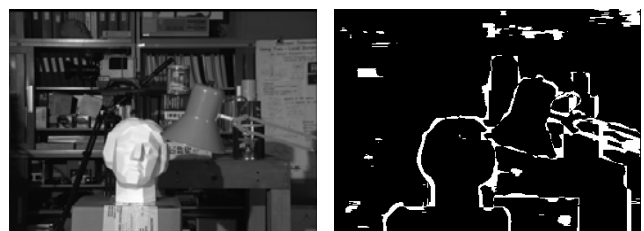


Fig. 10 – Imagem esquerda (esq.) e Mapa de pares errados para Tsukuba (dir.).

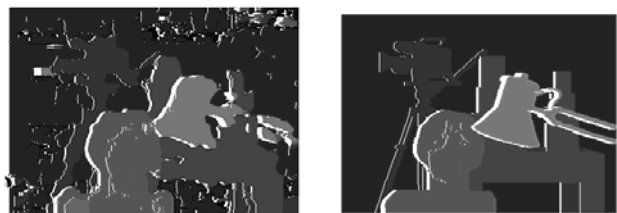


Fig. 11 – Mapas de disparidades esquerdos com oclusões reconstruído (esq.) e verdadeiro (dir.) para Tsukuba.

TABELA 3 – RESULTADOS PARA “TSUKUBA”.

	Tamanho da janela inicial						
	1x1	3x3	5x5	7x7	9x9	11x11	13x13
E(%)	23.1	14.0	10.1	7.9	6.8	6.4	6.7
ET(%)	33.9	25.1	17.3	12.2	9.1	7.6	7.1
EO(%)	37.6	39.8	47.8	56.8	65.9	72.8	77.8
T(s)	66	14	10	11	12	13	15

9.4 Imagens Venus

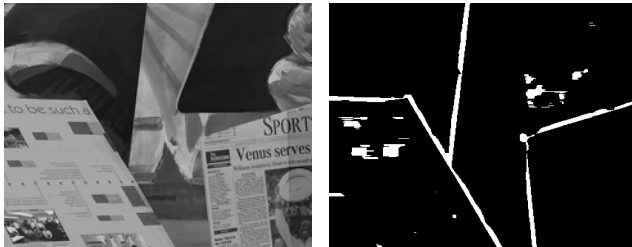


Fig. 12 – Imagem esquerda (esq.) e Mapa de pares errados para Vênus (dir.).

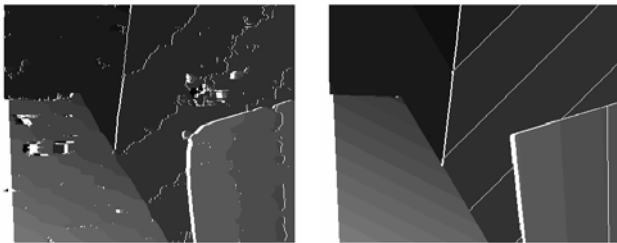


Fig. 13 – Mapas de disparidades esquerdos com oclusões reconstruído (esq.) e verdadeiro (dir.) para Venus.

TABELA 4 – RESULTADOS PARA “VENUS”.

	Tamanho da janela inicial						
	1x1	3x3	5x5	7x7	9x9	11x11	13x13
E(%)	27.1	14.6	8.8	5.9	4.7	4.2	4.0
ET(%)	40.2	30.2	18.1	11.1	7.8	6.5	5.7
EO(%)	23.9	21.2	22.2	23.0	24.8	25.6	25.5
T(s)	170	34	21	20	22	26	28

A figura 14 mostra um gráfico dos resultados (somatória dos erros E, ET e EO) disponíveis no *site* para as abordagens que não utilizam segmentação das imagens estereoscópicas e/ou dos mapas de disparidade a fim de determinar disparidades para as regiões segmentadas e que se comparam com a abordagem proposta que utiliza apenas os valores de intensidade das imagens estereoscópicas como fonte de informação para o processo estéreo.

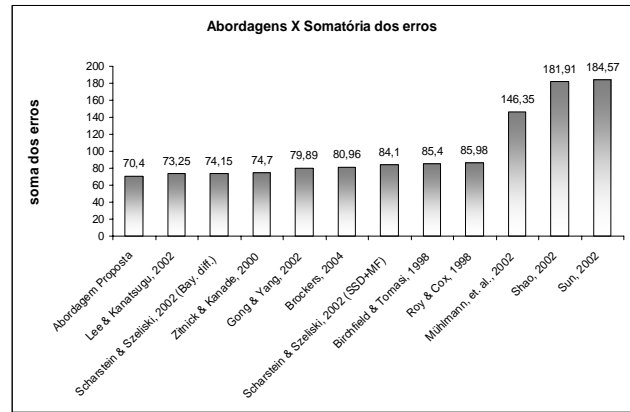


Fig. 14 – Gráfico das abordagens existentes versus a somatória dos erros.

A abordagem proposta foi testada também para Estereogramas de Pontos Aleatórios.

Estereogramas de Pontos Aleatórios podem ser gerados através do deslocamento relativo (disparidade) de regiões entre um par de imagens. Estas regiões homólogas recebem texturas aleatórias, porém iguais, de tal forma que a textura de uma região na imagem da esquerda é copiada para a região homóloga na imagem da direita.

Estereogramas de Pontos Aleatórios são importantes para avaliar a capacidade do algoritmo em resolver o Problema de Correspondência isento de considerações sobre intensidades e ruídos, de tal forma que a solução é derivada apenas em função da estrutura tridimensional da cena.

Os algoritmos que realizam algum tipo de segmentação a priori das imagens estereoscópicas, para posteriormente resolver o Problema de Correspondência, não conseguem determinar a estrutura tridimensional da cena para este tipo de imagens, uma vez que não existe nenhuma informação a respeito de formas, objetos ou outras feições de mais alto nível.

A figuras 15 mostra um par de Estereogramas de Pontos Aleatórios, as figuras 16 e 17 mostram os mapas de disparidade reconstruídos para as imagens da esquerda e da direita.

Em ambos os pares, o tamanho inicial da janela utilizado é $s_1^i \times s_1^j = 1 \times 1$ e 5 níveis hierárquicos.

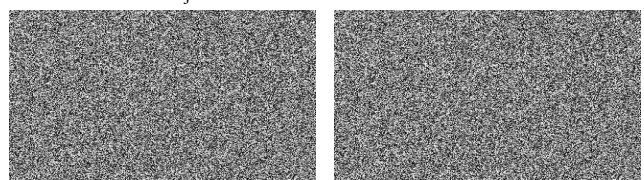


Fig. 15 – Exemplo de Estereograma de Pontos Aleatórios esquerdo e direito

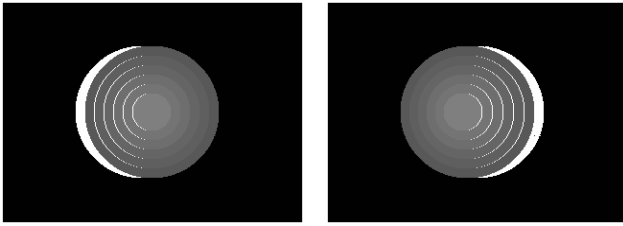


Fig. 16 – Mapas de disparidades com oclusões reconstruído para a imagem da esquerda (esq.) e da direita (dir.) para Redondo.

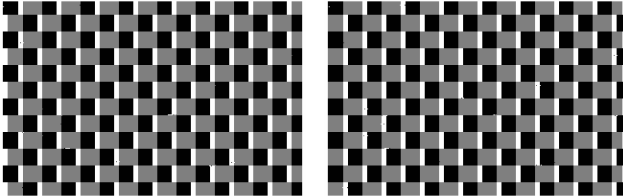


Fig. 17 – Mapas de disparidades com oclusões reconstruído para a imagem da esquerda (esq.) e da direita (dir.) para Xadrez.

A abordagem proposta reconstruiu os mapas de disparidades para os Estereogramas de Pontos Aleatórios com alta precisão (a porcentagem de pares errôneos é inferior a 0,01%).

10. CONCLUSÕES

De acordo com os resultados apresentados percebe-se a dependência dos resultados e do tempo de processamento em relação ao tamanho da janela inicial do processo hierárquico. Observa-se de modo geral que o erro e o tempo de processamento decrescem com o aumento da janela inicial até atingir o valor mínimo, voltando a crescer com o aumento do tamanho da janela inicial. Observa-se que não há coincidência dos pontos de mínimo para o erro e tempo de processamento.

No caso limite de janela inicial 1x1 e imagens com ruídos e texturas homogêneas se observa grande número de pares ambíguos, causando o aumento na porcentagem de pares errôneos e significativo aumento do tempo de processamento. No entanto, quando esses fatores (ruídos e texturas homogêneas) não são presentes, como nos Estereogramas de Pontos Aleatórios, o tamanho da janela inicial 1x1 apresenta os melhores resultados.

Pode-se afirmar que a utilização de janelas iniciais pequenas no processo de Correspondência Bidirecional Hierárquica preserva as descontinuidades e, conseqüentemente, melhora a detecção de oclusões. Em contrapartida, as regiões de superfícies suaves da cena não são bem reconstruídas devido ao aumento de pares ambíguos nestas regiões, principalmente nos casos de superfícies com texturas homogêneas e ruídos. Com isso pode-se concluir que o tamanho inicial ótimo da janela deve ser obtido em função de medidas que quantifiquem coerentemente a energia do processo ruidoso e das regiões com texturas homogêneas.

A identificação das regiões de oclusão do par de imagens estereoscópicas é de igual importância à

determinação dos valores de disparidade e, portanto, devem ser consideradas em qualquer algoritmo estéreo. Neste trabalho foi apresentada uma abordagem consistente para a detecção das regiões de oclusões através da imposição da Injunção de Ordem.

Quanto à comparação dos resultados com outras abordagens pode-se concluir que a abordagem proposta foi superior às demais abordagens que não são baseadas em Segmentação e/ou se utilizam de otimização de funcionais cujos parâmetros são determinados de maneira empírica. Os resultados alcançados com a abordagem proposta são, em geral, inferiores aos resultados alcançados pelas abordagens recentes baseadas em Segmentação, porém, como já citado, tais abordagens utilizam um grande número de parâmetros definidos de maneira empírica, os quais dependem das condições de iluminação e principalmente da própria estrutura tridimensional da cena. No caso da abordagem proposta somente o tamanho da janela inicial deve ser definido e para este parâmetro verifica-se que a utilização de uma janela de tamanho maior que o ótimo não resulta em erros ou tempo de processamento significativamente maiores, tornando a abordagem flexível em relação à definição dos parâmetros iniciais e praticamente independente destes no caso de maior tolerância em relação ao erro.

REFERÊNCIAS BIBLIOGRÁFICAS

- BELHUMEUR, P.N. A binocular stereo algorithm for reconstructing sloping, creased, and broken surfaces in the presence of half-occlusion. Proceedings of Fourth International Conference on Computer Vision, 1993. Pag: 431-438.
- BLEYER, M. & GELAUTZ, M. Graph-based surface reconstruction from stereo pairs using image segmentation. Proceedings of SPIE, vol. 5665, January 2005. Pag: 288-299.
- BOYKOV, Y.; VEKSLER, O. & ZABIH, R. Fast approximate energy minimization via graph cuts. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2001. 23(11):1222-1239.
- DARRELL, T. & PENTLAND, A. Cooperative robust estimation using layers of support. IEEE Transactions on Pattern Analysis and Machine Intelligence 1995, 17(5):474-487.
- DENG, Y.; YANG, Q., LIN, X. & TANG, X. A symmetric patch-based correspondence model for occlusion handling. IEEE Transactions on Pattern Analysis and Machine Intelligence 2007, 29(6): 1068-1079.
- HORN, B.K.P. Robot Vision. MIT Press, Cambridge, MA, 1985.

ISHIKAWA, H. & GEIGER, D. Occlusions, discontinuities, and epipolar lines in stereo. Proceedings of the European Conference on Computer Vision, volume 1, pag. 232-249, 1998.

KOLMOGOROV, V. & ZABIH, R. Computing visual correspondence with occlusions using graph cuts. IEEE International Conference on Computer Vision, 2001. Pag: 508-515.

LIN, M. & TOMASI, C. Surfaces with occlusions from layered stereo. Ph.D. thesis, Stanford University, 2002.

NOGUEIRA, F.M.A. & TOZZI, C.L. Geração Automática de Mapas de Disparidade em Visão Estéreo. Dissertação de Mestrado. FEEC/UNICAMP, Campinas, 1998.

OHTA, Y. & KANADE, T. Stereo by intra- and inter-scanline search using dynamic programming. IEEE Transactions on Pattern Analysis and Machine Intelligence, 7(2):139-154, 1985.

POGGIO, T.; TORRE, V. & KOCH, C. Computational vision and regularization theory. Nature, 317:314-319, 1985.

POLLARD, S. B.; PORRILL, J.; J. MAYHEW, E.W., & FRISBY, J. P. Disparity gradient, Lipschitz continuity and computing binocular correspondences. Proceedings of 3rd Int. Symposium of Robotics Research, 1986, PAG: 19-26.

ROY, S. & COX, I. J. A maximum-flow formulation of the N-camera stereo correspondence problem. IEEE International Conference on Computer Vision, 1998. PAG: 492-499.

SCHARSTEIN, D. & SZELISKI, R. A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms, International Journal of Computer Vision, 47(1-3) 7-42, 2002.

TERZOPOULOS, D. Regularization of inverse visual problems involving discontinuities. IEEE Transactions on Pattern Analysis and Machine Intelligence, 8(4):413-424, 1986.

ZITNICK, L. & KANADE, T. A cooperative algorithm for stereo matching and occlusion detection. IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(7):675-684, 2000.