



Ingeniare. Revista Chilena de Ingeniería

ISSN: 0718-3291

facing@uta.cl

Universidad de Tarapacá

Chile

Pasquali, André Luiz; Romero Romero, Milton Ernesto; Mazina Martins, Evandro
Mapa denso de disparidade para imagem estereoscópica no domínio de Clifford
Ingeniare. Revista Chilena de Ingeniería, vol. 21, núm. 3, diciembre, 2013, pp. 424-437
Universidad de Tarapacá
Arica, Chile

Disponível em: <http://www.redalyc.org/articulo.oa?id=77228820012>

- Como citar este artigo
- Número completo
- Mais artigos
- Home da revista no Redalyc

Mapa denso de disparidade para imagem estereoscópica no domínio de Clifford

Dense disparity map for the color stereoscopic image on the Clifford domain

André Luiz Pasquali¹ Milton Ernesto Romero Romero¹ Evandro Mazina Martins¹

Recibido 29 de mayo de 2012, aceptado 20 de mayo de 2013

Received: May 29, 2012 Accepted: May 20, 2013

RESUMO

Extensões do *Quad-tree* para a estimação do mapa denso de disparidade utilizando a luminância no domínio real foram propostas na literatura. A contribuição deste trabalho é a comparação do processamento de imagens estéreo entre o algoritmo que é uma extensão do *Quad-tree* para processar imagens no domínio real utilizando a luminância, com o algoritmo, aqui proposto, que é a extensão do *Quad-tree* para processar imagens codificadas no domínio da álgebra de *Clifford* utilizando imagens coloridas do modelo YUV. O mapa denso de disparidade é obtido resolvendo o problema da correspondência para cada *pixel* entre o par estereoscópico. O problema da correspondência é fundamental e apresenta dificuldades devido à presença de áreas oclusas, ruídos e variações na iluminação, etc. A álgebra de *Clifford* é utilizada para mapear os valores do modelo YUV de cada *pixel* nos vetores de base, compondo um número de *Clifford*. Neste domínio, geram-se resultados melhores para imagens que têm muitos objetos na cena, aqui definidas como complexas, se comparados com os obtidos sobre as imagens utilizando o domínio real sobre a luminância. Processando os pares estereoscópicos utilizados no algoritmo proposto no domínio real e no domínio de *Clifford* se demonstra, experimentalmente, a melhora obtida para imagens em cenas reais.

Palavras chave: *Quad-tree*, álgebras de *Clifford*, estereoscopia, mapa da disparidade, YUV.

ABSTRACT

Extensions of the Quad-tree for the estimation of dense disparity map using the luminance on the real domain have been proposed in the literature. Main focus of this work is the comparison between the algorithms for the processing of stereoscopic images utilizing the luminance Y in the real domain by an extension of the Quad-tree, with the extended Quad-tree, here proposed, to process color images encoded in the field of Clifford algebras for the YUV model. The dense disparity map is obtained by solving the correspondence of pixels for each pixel of the stereoscopic pair. The correspondence problem is an essential task that presents problems due to the presence of occluded areas, noise and variations in lighting, among others. The Clifford algebra is used to map the color values YUV of each pixel on the basis vectors, thus making a single Clifford number. In this domain, better results for images with many objects in the scene are generated, as compared with those obtained on the images using the luminance on the real domain. By processing multiple stereoscopic pairs, for both algorithms, the improvement obtained in real scenes is experimentally demonstrated.

Keywords: *Quad-tree*, *Clifford algebra*, *stereoscopic processing*, *disparity map*, *YUV*.

¹ Departamento de Engenharia Elétrica. Universidade Federal de Mato Grosso do Sul. Caixa Postal 549, CEP 79070-900. Campo Grande-MS, Brasil. E-mail: pasqualiec@gmail.com; miltonr@del.ufms.br; evandro.mazina@ufms.br

INTRODUÇÃO

A estereoscopia possibilita a reconstrução de estruturas tridimensionais através de duas imagens de uma mesma cena, mas de ângulos ligeiramente defasados, como mostram as imagens esquerda e direita do par estereoscópico da Figura 1.

Os *pixels* correspondentes são aqueles que dado um *pixel* de um objeto na cena, esteja presente tanto na imagem esquerda quanto na direita [1]. Essas reconstruções são possíveis porque utilizando o par estereoscópico se pode estimar o deslocamento ou disparidade de cada *pixel* pertencente à imagem esquerda em relação à direita ou vice versa, isto é conhecido como o problema da correspondência. A estimativa do mapa denso de disparidade (MDD) consiste em encontrar a correspondência para todos os *pixels* da imagem esquerda em relação à direita ou vice versa, e apresenta problemas causados por regiões oclusas, ruídos e variações na iluminação das imagens utilizadas, diferentes ganhos das câmeras, etc. A visão tridimensional (3D) é utilizada em aplicações que envolvem visão em profundidade, como por exemplo: a visão de robôs, a realidade virtual, a realidade virtualizada, a TV 3D [2], etc.

Várias abordagens têm sido utilizadas para resolver o problema da correspondência utilizando a correlação entre blocos [3-4], a otimização de gradiente [3], a correspondência de características [5], o *Quad-tree* [2], análise multiescala [6-8], entre outras [9-11].

O algoritmo *Quad-tree*, de análise multiresolução, partitiona a imagem em quatro quadrantes ou regiões aplicando filtros e decimando de maneira recursiva, para gerar os ícones à baixa resolução. Para o caso da imagem estereoscópica, no algoritmo proposto em Chiari, Ciciani, Romero e Rossi [2], nos ícones de menor resolução se considera unicamente a região de mais baixa frequência para gerar a primeira estimativa do mapa de disparidade, as outras regiões que seriam geradas pelo *Quad-tree* não são utilizadas. Posteriormente, se aumenta, de forma recursiva, a resolução por meio de *upsampling* para gerar o mapa de disparidade a plena resolução.

A contribuição neste trabalho é a comparação do processamento de imagens estéreo entre o algoritmo apresentado por Chiari, Ciciani, Romero e Rossi [2], que é uma extensão do *Quad-tree* para processar



(a)



(b)

Figura 1. Par estereoscópico imagens do Trem. (a) imagem esquerda, (b) imagem direita.

imagens no domínio real utilizando a luminância, com o algoritmo proposto neste trabalho, que é a extensão do *Quad-tree* para processar imagens codificadas no domínio da álgebra de *Clifford* [9, 10, 12, 13] utilizando imagens coloridas no modelo YUV. O domínio de *Clifford* utilizado neste trabalho, isomorfo ao espaço Euclidiano 3D, é composto por um escalar 1; 3 vetores e_1, e_2, e_3 ; 3 bivetores e_{12}, e_{13}, e_{23} ; e o pseudoescalar e_{123} para a assinatura $Cl_{3,0}$. O escalar é uma grandeza adimensional ou número puro. Os vetores no domínio de *Clifford* são análogos ao i, j, k do espaço Euclidiano, os bivetores são áreas orientadas e o pseudoescalar ou trivector é um volume orientado.

Para trabalhar com imagens coloridas em *Clifford*, se realiza um mapeamento entre os valores de luminância Y, e as crominâncias U e V de cada *pixel* nos vetores da álgebra.

As imagens usadas no cálculo do mapa denso de disparidade são as mesmas utilizadas no trabalho em *Chiari, Ciciani, Romero e Rossi* [2] para poder comparar os resultados reportados do processamento sobre a luminância, com o do processamento a cores do modelo YUV, no domínio de *Clifford*. Cada imagem do par estéreo tem resolução de 720 por 576 pixels, utilizada na TV 3D, sendo 25 pares por segundo, no sistema europeu.

Os resultados do *Quad-tree* estendido no domínio de *Clifford* aqui proposto, comparados com os obtidos em [2], mostram tanto uma maior relação S/N (*Signal-to-Noise Ratio*), quanto um melhor resultado visível sobre as imagens complexas. Imagens complexas são aqui definidas como aquelas que têm mais objetos presentes na cena a diferentes profundidades. Nas imagens simples (poucos objetos) foram obtidos resultados semelhantes para os dois algoritmos.

O restante deste artigo está organizado da seguinte forma: a seguir é apresentada a fundamentação teórica composta pela estereoscopia, seguida da descrição detalhada do *Quad-tree* proposto em [2], e continuando com os conceitos fundamentais da álgebra de *Clifford*. Na seguinte seção é apresentada a extensão aqui proposta do *Quad-tree* na álgebra de *Clifford*. São abordados, também, os resultados com a discussão, seguido das conclusões e trabalhos futuros.

FUNDAMENTAÇÃO TEÓRICA

Estereoscopia

A estereoscopia possibilita a reconstrução de estruturas tridimensionais utilizando duas imagens de diferentes perspectivas (par estereoscópico) como mostradas na Figura 1. Essas reconstruções são realizadas por meio do mapa de disparidade que possui informações sobre o deslocamento dos pixels correspondentes do par estereoscópico. A qualidade das reconstruções depende da qualidade da estimativa do mapa de disparidade dos pixels correspondentes entre as imagens.

Encontrar o pixel correspondente entre duas imagens estéreo é uma tarefa fundamental, pois o pixel pode estar presente em apenas uma das imagens, caso ele pertença a uma área oclusa ou tenha sofrido alguma interferência causada pela diferença de iluminação entre as imagens estereoscópicas, e até

mesmo em casos onde o dispositivo de captura das imagens está descalibrado. Técnicas de suavização e a aplicação da restrição de geometria epipolar, sem deslocamento vertical [3], são comumente usadas para amenizar esses problemas.

Os pixels correspondentes são estimados por meio de uma função de correlação que determina quanto dois pixels ou duas regiões são semelhantes. Existem diversas funções de correlação presentes na literatura como: *Normalized Cross-Correlation* (NCC), *Sum of Squared Differences* (SSD), *Normalized SSD*, etc. [3-5] e [14-15]. Neste trabalho aplica-se o *Mean Absolute Difference* (MAD) dada pela equação (1), para o cálculo do pixel na posição (x, y) denotado $P_R(x,y)$ da imagem de referência, correspondente na outra imagem ($P_L(x,y)$). Onde S_h, S_v são os limites da área de busca e B_h, B_v , são os limites do bloco horizontal e vertical, respectivamente.

$$\begin{aligned} MAD(m,n) = \\ \sum_{h} \sum_{k} |[P_R(x+h, y+k) - P_L(x+m+h, y+n+k)]| \\ h = -B_h, \dots, B_h; k = -B_v, \dots, B_v \\ m = -S_h, \dots, S_h; n = -S_v, \dots, S_v; \end{aligned} \quad (1)$$

Na equação (1), o par (x,y) corresponde, para o exemplo da Figura 2, às coordenadas do pixel na imagem, o par (m,n) corresponde à área de busca (em vermelho), o par (h,k) corresponde ao tamanho do bloco (em azul). Utilizando a técnica conhecida como busca completa (*Block Matching*) [2] como mostra a Figura 2, se estima o mapa denso de disparidade para todos os pontos na imagem de referência. Nesta figura cada pixel é representado por um quadrado e o quadrado em consideração é o $P_R(x,y)$ (em verde no ícone direito).

O Mapa Denso de Disparidade (MDD) se estima de acordo com a equação (2) sendo a disparidade

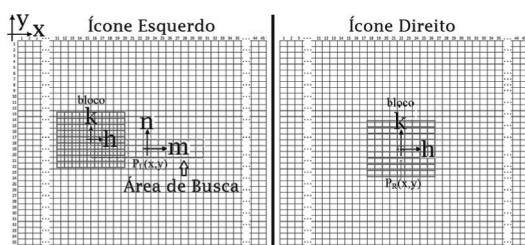


Figura 2. Exemplo de área de busca completa.

horizontal o valor do índice m e a disparidade vertical o valor do índice n para o valor da $MAD(m,n)$ mínimo entre os valores considerados na área de busca.

$$MDD(x,y) = (m,n) \text{ para } \min[MAD(m,n)] \quad (2)$$

A disparidade pode ser positiva, negativa ou nula. Como mostra a Figura 3, a disparidade é positiva quando o *pixel* da imagem esquerda é projetado no plano de projeção à direita do seu *pixel* correspondente da imagem direita.

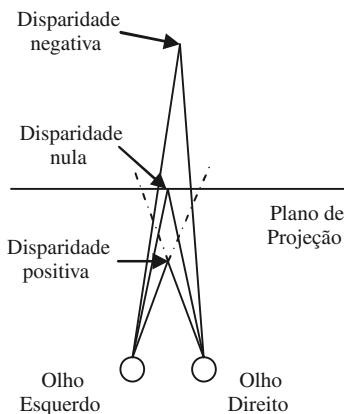


Figura 3. Plano de projeção da disparidade.

A disparidade é negativa quando o *pixel* da imagem esquerda é projetado à esquerda do seu *pixel* correspondente da imagem direita. A disparidade é nula quando o *pixel* da imagem direita e seu *pixel* correspondente da imagem esquerda estão projetados no mesmo ponto do plano de projeção.

No mapa denso de disparidades pontos que possuem disparidade positiva são representados por tons de cinza mais claro; os pontos com disparidade negativa são representados por tons de cinza mais escuro; e os pontos com disparidade nula são representados por um tom de cinza mediano; onde objetos mais próximos do observador ficam representados em tons claros e, portanto, o mapa de disparidade é um mapa da profundidade dos objetos na imagem, como mostra a Figura 4.

QUAD-TREE

Em [2] o algoritmo *Quad-tree* de análise multirresolução, toma cada imagem, esquerda e direita, à plena resolução de 720 por 576 para o sistema de TV europeu e aplica uma convolução

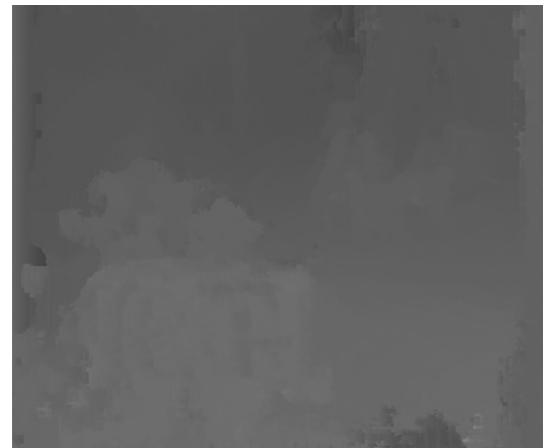


Figura 4. Mapa denso de disparidade do par estereoscópico da Figura 1.

com o filtro *antialiasing* seguida de uma decimação por 2 gerando dois novos ícones na resolução de 360 por 288. Tanto a filtragem quanto a decimação são repetidas sempre sobre a imagem de mais baixa frequência unicamente (note que o *Quad-tree* particiona a imagem em quatro quadrantes ou regiões com diferentes faixas de frequências, somente se utiliza a de mais baixa frequência para esta proposta), até que sejam produzidos dois ícones com resolução de $x=45$ por $y=36$, um referente à imagem direita e outro à esquerda. Com o par de ícones esquerdo e direito os valores utilizados para calcular o MAD como mostra a equação (1) para o algoritmo proposto, serão: $x=[1,45]$, $y=[1,36]$; $m \in [-7, 7]$ e $n \in [-1, 1]$, e tamanho do bloco $h \in [-4,4]$ e $k \in [-4,4]$. Para este suporte de 9 por 9 *pixels* centrados sob o *pixel* em consideração é estimada a disparidade de cada *pixel* gerando assim o primeiro mapa denso de disparidade para essa resolução. Posteriormente, se aumenta a resolução multiplicando por 2 da disparidade do mapa de 45 por 36 e se duplica a posição gerando um novo mapa com valores conhecidos nas posições pares e valores desconhecidos nas ímpares. O processo seguinte é atualizar e interpolar esses valores para completar o mapa denso de disparidade na resolução de 90 por 72. Esse processo se repete até se obter o mapa a plena resolução de 720 por 576.

Na Figura 5, a extensão proposta em Chiari, Ciciani, Romero e Rossi [2] utiliza 5 níveis de resolução para estimar a disparidade de um par de imagens estereoscópicas. Este algoritmo possui menor

complexidade computacional se comparado a outros algoritmos de busca completa que são utilizados para o cálculo de disparidade das imagens a plena resolução.

Tabela 1. Coeficientes do filtro *antialiasing*.

Coeficientes	Valor	Coeficientes	Valor
h0	-0,005615	h13	0,314697
h1	-0,005127	h14	0,029785
h2	0,011230	h15	-0,095703
h3	0,012451	h16	-0,025635
h4	-0,017090	h17	0,047119
h5	-0,023193	h18	0,021484
h6	0,021484	h19	-0,023193
h7	0,047119	h20	-0,017090
h8	-0,025635	h21	0,012451
h9	-0,095703	h22	0,011230
h10	0,029785	h23	-0,005127
h11	0,314697	h24	-0,005615
h12	0,471191		

A estrutura do *Quad-tree*, apresentada na Figura 5, é descrita a seguir:

1. Balancear a luminância das imagens de entrada, onde é aplicado um filtro de suavização para compensar possíveis desajustes do dispositivo de captura das imagens, onde L e R representam a luminância da imagem esquerda e direita, respectivamente.

2. Aplicar o filtro *antialiasing*, onde é realizada uma convolução entre a imagem e o filtro presente na Tabela 1. Nesta tabela se mostram os 25 coeficientes do filtro passa baixa, e na Figura 6 se apresentam os mesmos 25 coeficientes no domínio do tempo no eixo horizontal e o valor do coeficiente no eixo vertical do filtro utilizado no processo. Como o filtro utilizado é linearmente separável e a imagem é bidimensional, o filtro pode ser aplicado em dois passos, inicialmente na horizontal e posteriormente na vertical.

O filtro LPF_2, da Tabela 2, produz uma versão filtrada da imagem original, que pode ser decimada por 2 na horizontal e vertical sem a introdução de *aliasing* na imagem resultante.

3. *Downsampling*, decimar a imagem por 2. Nesse passo é criado um ícone, com metade da resolução

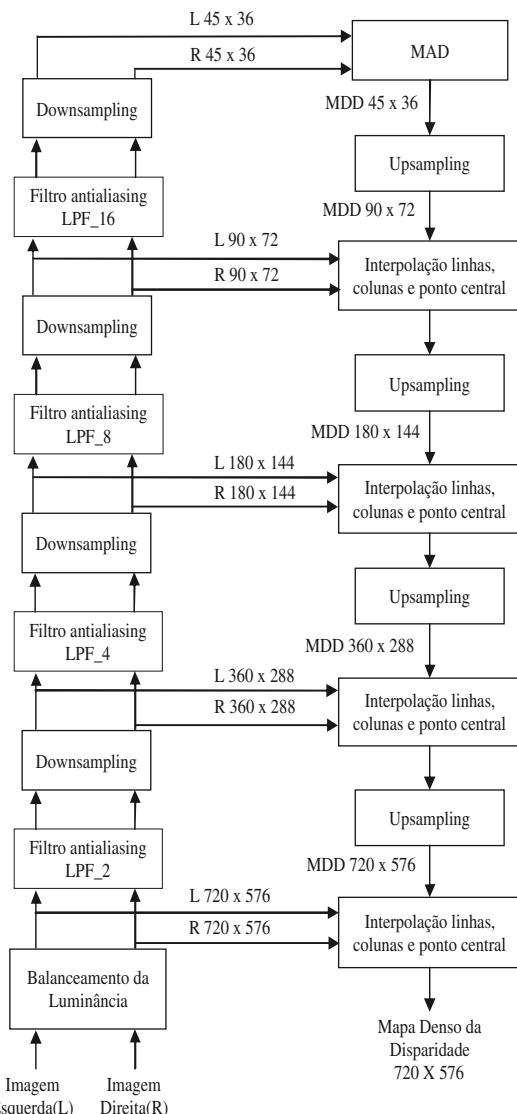


Figura 5. Estrutura do *Quad-tree*.

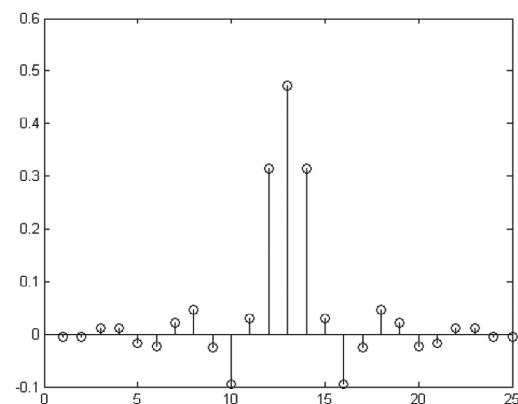


Figura 6. Filtro *antialiasing* (Tabela 1).

da imagem filtrada, formado apenas com os *pixels* pares da imagem.

4. Repetir os passos 2 e 3 até que a imagem tenha a resolução de 45 por 36.

5. Calcular o mapa denso da disparidade, MDD como mostra a equação (2), utilizando o MAD, como mostra a equação (1). Com o ícone de 45 por 36 é criado o primeiro MDD também de 45 por 36 que será utilizado na fase de *upsampling*.

6. *Upsampling* por 2 na resolução e multiplicar por 2 os valores do MDD. A Figura 7 mostra como é realizada a expansão de cada *pixel*, onde os quadrados representam os pontos com disparidade conhecida e os círculos os pontos com disparidade desconhecida. Observando o bloco de 2 por 2 do nível K, o valor de disparidade presente na posição $j = 8, i = 2$ é multiplicado por dois e armazenado na posição $j = 16, i = 4$ no nível $K+1$. Para calcular o MDD é necessária a posição do *pixel* (ponto) em consideração, e centrado neste *pixel*, se define o bloco N por N com $N=9$ no ícone de luminância onde se encontra a área de busca para o cálculo do MAD.

7. Atualização de disparidade das posições pares do MDD, passo 1, Figura 8.

Tabela 2. Resposta impulso filtro LPF.

Filtro	Função de Filtragem
LPF_2	$h_2[n] = h[n]$
LPF_4	$h_4[n] = h[n] * h[n]$
LPF_8	$h_8[n] = h[n] * h[n] * h[n]$
LPF_16	$h_{16}[n] = h[n] * h[n] * h[n] * h[n]$

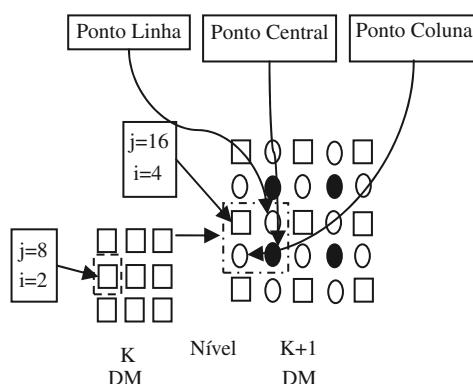


Figura 7. Expansão de cada pixel do mapa de disparidade do nível K ao K+1 (*Upsampling*).

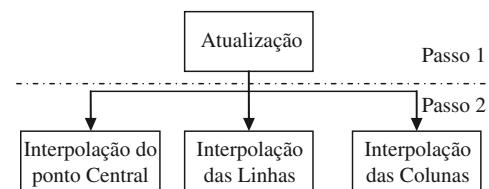


Figura 8. Sequência de passos para atualização e interpolação do mapa de disparidade.

A seguir se explica como se escolhe a área de busca para executar o processo de expansão mostrado na Figura 7.

A área de busca para atualização da disparidade são pontos formados pelos vizinhos ao ponto em consideração ($i=4, j=16$, nível $K+1$). Utilizando a disparidade da posição ($i=4, j=16$) e com os ícones de luminância direito e esquerdo com resolução de 90 por 72 é atualizada a disparidade aplicando a função de correlação MAD sobre a vizinhança ao *pixel* sob consideração, representado pelo quadrado ($i=4, j=16$) para uma área de busca de 3 por 3 formada pelos *pixels* vizinhos, veja a Figura 7.

8. A interpolação do ponto central, das linhas e colunas, passo 2 na Figura 8, é realizada na seguinte ordem:

- (a) A disparidade do ponto central (círculos escuros na Figura 7) é estimada em dois passos: primeiro é calculado o baricentro dos seus quatro vizinhos conhecidos (quadrados em torno do círculo escuro na Figura 7), e depois a disparidade do ponto central recebe o índice de menor MAD dos 5 pontos (área de busca): o baricentro e os 4 vizinhos conhecidos;
- (b) Interolar o ponto central, as linhas e as colunas. A disparidade dos pontos linhas (círculos brancos sinalizados na Figura 7) é estimada utilizando a média do valor de disparidade de seus vizinhos localizados a esquerda e a direita. Com o valor da média e de disparidade dos vizinhos localizados a esquerda e a direita é calculado o MAD para estes 3 pontos (área de busca) e o menor valor resultante é armazenado no ponto linha. A disparidade dos pontos colunas (círculos brancos sinalizados na Figura 7) é calculada utilizando a média dos vizinhos localizados acima e abaixo. Com o valor da média e de

disparidade dos vizinhos localizados acima e abaixo é calculado o MAD sob cada um dos três valores (área de busca) e o menor valor de disparidade retornado pelo MAD é armazenado no ponto coluna.

9. Repetir os passos 6, 7 e 8 até que o MDD tenha a resolução de 720 por 576.

ÁLGEBRA DE CLIFFORD

A álgebra de *Clifford* é um sistema adaptado à geometria ortogonal de um espaço arbitrário, que de acordo com sua assinatura pode generalizar diversos espaços. A assinatura é representada por $\text{Cl}_{p,q} = \text{Cl}_n$ onde:

- p : é igual à quantidade de vetores de base que elevados ao quadrado são iguais a 1;
- q : é igual à quantidade de vetores de base que elevados ao quadrado são iguais a -1;
- $n = p + q$: é a soma da quantidade de p e q .

O domínio de *Clifford* isomorfo ao espaço Euclidiano 3D utilizado neste trabalho é a álgebra de *Clifford* de assinatura $\text{Cl}_{3,0} = \text{Cl}_3$ ($p=n=3$, $q=0$).

O produto de *Clifford* entre vetores, representado pelo símbolo “*” é definido na Tabela 3 para a assinatura Cl_3 . O produto é a combinação linear da parte simétrica: o produto interior representado pelo símbolo “.” e da parte antissimétrica: o produto exterior representado pelo símbolo “^”, como mostra a equação (3). Estes produtos estão definidos nas Tabelas 4, 5, respectivamente. Os elementos gerados pelo produto exterior (por exemplo: $e_1 \wedge e_2 \wedge \dots \wedge e_k$) são chamados de *k-blade*, como o escalar 1 (0 *blade*); os vetores e_1, e_2, e_3 (1 *blade*); os bivetores e_{12}, e_{13}, e_{23} (2 *blade*); e o pseudoescalar e_{123} (3 *blade*). As combinações lineares de *k-blades* são chamadas de multivetores ou números de *Clifford*. As equações (3), (4) e (5) mostram as relações entre o produto de *Clifford* conhecendo os produtos interno e externo e vice-versa.

$$a^* b = a \cdot b + a \wedge b \quad (3)$$

$$a \cdot b = \frac{1}{2}(a^* b + b^* a) = b \cdot a \quad (4)$$

$$a \wedge b = \frac{1}{2}(a^* b - b^* a) = -b \wedge a \quad (5)$$

Tabela 3. Produto de *Clifford*.

Cl_3	1	e_1	e_2	e_3	e_{12}	e_{31}	e_{23}	e_{123}
1	1	e_1	e_2	e_3	e_{12}	e_{31}	e_{23}	e_{123}
e_1	e_1	1	e_{12}	$-e_{31}$	e_2	$-e_3$	e_{123}	e_{23}
e_2	e_2	$-e_{12}$	1	e_{23}	$-e_1$	e_{123}	e_3	e_{31}
e_3	e_3	e_{31}	$-e_{23}$	1	e_{123}	e_1	$-e_2$	e_{12}
e_{12}	e_{12}	$-e_2$	e_1	e_{123}	-1	e_{23}	$-e_{31}$	e_3
e_{31}	e_{31}	e_3	e_{123}	$-e_1$	$-e_{23}$	-1	e_{12}	e_2
e_{23}	e_{23}	e_{123}	$-e_3$	e_2	e_{31}	$-e_{12}$	-1	e_1
e_{123}	e_{123}	e_{23}	e_{31}	e_{12}	$-e_3$	$-e_2$	$-e_1$	-1

A adição e subtração em álgebra de *Clifford* são análogas à adição e subtração dos números complexos, onde todos esses elementos podem ser somados e subtraídos em cada dimensão.

Tabela 4. Produto interior.

Cl_3	1	e_1	e_2	e_3	e_{12}	e_{31}	e_{23}	e_{123}
1	1	e_1	e_2	e_3	e_{12}	e_{31}	e_{23}	e_{123}
e_1	0	1	0	0	e_2	$-e_3$	0	e_{23}
e_2	0	0	1	0	$-e_1$	0	e_3	e_{31}
e_3	0	0	0	1	0	e_1	$-e_2$	e_{12}
e_{12}	0	0	0	0	-1	0	0	$-e_3$
e_{31}	0	0	0	0	0	-1	0	$-e_2$
e_{23}	0	0	0	0	0	0	-1	$-e_1$
e_{123}	0	0	0	0	0	0	0	-1

Detalhes sobre o produto geométrico, interior e exterior encontram-se em [12-13].

A convolução P(m,n) em *Clifford* [16], análoga a convolução no domínio real, é realizada por meio do produto de *Clifford* entre o filtro e a imagem, ambos definidos no domínio de *Clifford* como mostra a equação (6), onde $(-X, X)$ e $(-Y, Y)$ são os limites para o cálculo da convolução da imagem com o filtro.

Tabela 5. Produto exterior.

Cl_3	1	e_1	e_2	e_3	e_{12}	e_{31}	e_{23}	e_{123}
1	1	e_1	e_2	e_3	e_{12}	e_{31}	e_{23}	e_{123}
e_1	e_1	0	e_{12}	$-e_{31}$	0	0	e_{123}	0
e_2	e_2	$-e_{12}$	0	e_{23}	0	e_{123}	0	0
e_3	e_3	e_{31}	$-e_{23}$	0	e_{123}	0	0	0
e_{12}	e_{12}	0	0	e_{123}	0	0	0	0
e_{31}	e_{31}	0	e_{123}	0	0	0	0	0
e_{23}	e_{23}	e_{123}	0	0	0	0	0	0
e_{123}	e_{123}	0	0	0	0	0	0	0

$$P(m,n) = \sum_{x=-X}^X \sum_{y=-Y}^Y P(m,n) * h(m-x, n-y) \quad (6)$$

Para o cálculo do MAD definido no domínio de *Clifford* (MADC) utiliza-se a norma dos multivetores para a assinatura $p=3$ utilizada neste trabalho, que é calculada de forma similar ao valor absoluto aplicado sobre o conjunto dos números complexos, ou seja, sendo A um multivetor, como mostra a equação (7), a norma de $|A|$ é a raiz quadrada da soma dos λ_i (escalares) elevados ao quadrado, como mostra a equação (8), e resulta em um único valor positivo que representa a grandeza do multivetor.

$$\begin{aligned} A = & \lambda_1 + \lambda_2 \cdot e_1 + \lambda_3 \cdot e_2 + \lambda_4 \cdot e_3 + \\ & \lambda_5 \cdot e_{12} + \lambda_6 \cdot e_{23} + \lambda_7 \cdot e_{31} + \lambda_8 \cdot e_{123} \end{aligned} \quad (7)$$

$$|A| = \sqrt{\lambda_1^2 + \lambda_2^2 + \lambda_3^2 + \lambda_4^2 + \lambda_5^2 + \lambda_6^2 + \lambda_7^2 + \lambda_8^2} \quad (8)$$

O MADC é definido sobre o domínio de *Clifford* como mostra a equação (9), note que a forma é a mesma da MAD no domínio real, mas os *pixels* estão representados no domínio de *Clifford* e a subtração e soma também são operadores no domínio de *Clifford*. Análogo ao domínio real, a equação (10) define o valor de disparidade como os índices (m, n) que correspondem ao mínimo obtido com o cálculo da MADC para as diferentes comparações do *pixel* de referência com os *pixels* na área de busca e, portanto, (m, n) são os deslocamentos (disparidade) horizontal e vertical, respectivamente.

$$\begin{aligned} MADC(m,n) = & \\ & \sum_h \sum_k \left[[P_R(x+h, y+k) - P_L(x+m+h, y+n+k)] \right] \\ & h = -Bh, \dots, Bh; k = -Bv, \dots, Bv \\ & m = -Sh, \dots, Sh; n = -Sv, \dots, Sv; \end{aligned} \quad (9)$$

$$MDD(x,y) = (m,n) \text{ para } \min |MADC(m,n)| \quad (10)$$

EXTENSÃO DO QUAD-TREE PARA O DOMÍNIO DE CLIFFORD

Neste trabalho o objetivo central é comparar o processamento utilizando um mapeamento entre os valores de luminância Y , e as crominâncias U e V de cada *pixel* nos vetores da álgebra (e_1, e_2, e_3), como mostra a equação (11), em relação ao processamento similar com os *pixels* da luminância mapeados sobre o domínio real. Entende-se por “similar” no sentido das adaptações aqui propostas para o domínio de *Clifford* que trabalha em YUV do algoritmo proposto por *Chiari, Ciciani, Romero e Rossi* [2] que trabalha apenas com a luminância definida no domínio dos reais.

$$P_{x,y} = Y_{x,y} * e_1 + U_{x,y} * e_2 + V_{x,y} * e_3 \quad (11)$$

Note que as componentes YUV são mapeadas nos vetores (e_1, e_2, e_3), cujos subscritos (x,y) são as coordenadas do *pixel* na imagem sob análise.

Utilizando a abordagem proposta em *Chiari, Ciciani, Romero e Rossi* [2] e a álgebra de *Clifford*, o algoritmo *Quad-tree* pode ser estendido para trabalhar com imagens coloridas. A extensão em *Clifford* segue a estrutura presente na Figura 5, onde foram realizadas as seguintes alterações necessárias para utilizar a álgebra de *Clifford*.

Para o passo 1, balanço na luminância, a imagem colorida é definida em *Clifford* [17-19] como descrito pela equação (11), onde cada *pixel* é representado por um multivetor.

Para o passo 2 o filtro com valores definidos na Tabela 1 e mostrados na Figura 6, é codificado no escalar 1, do domínio de *Clifford*. Neste passo se utiliza a convolução da equação (6), onde o símbolo “*” representa o produto de *Clifford* e as variáveis ($P(x,y), h(m-x, n-y)$) são números de *Clifford*. É importante notar que esta escolha implica que o filtro não é codificado no espaço dos bivetores ou do pseudoescalar porque aumentaria a complexidade computacional e a informação que seria obtida é redundante em relação às necessidades nesta aplicação. Note que a aplicação do filtro codificado no escalar tem como resultado filtrar cada um das componentes espectrais de cores com esse filtro passa-baixas.

Para o passo 3 o processo de decimação por 2 não tem alteração.

Para o passo 4 a repetição do processo de filtragem não tem alteração.

Para o passo 5 é utilizado o MADC, como mostra a equação (9), para determinar a correspondência dos *pixels*.

Para os passos 6, 7, 8 e 9 não têm alteração.

Note que no algoritmo *Quad-tree* esta sendo utilizada unicamente a imagem gerada de menor frequência (região gerada pelo filtro da banda mais baixa em frequência). A ideia fundamental é que o processo de decimação por 2 aplicado de maneira repetitiva gera um ícone com resolução 16 vezes menor (45 por 36) onde cada *pixel*, nesta resolução, “contém” informação referente a uma área 16 vezes maior a plena resolução (720 por 576). Como o bloco utilizado é de 9 por 9 para a aplicação da MAD, “equivale” a uma área de comparação a plena resolução de 144 *pixels* por 144 *pixels* para resolver o problema da correspondência. Esta é uma boa redução do problema da complexidade computacional que é da ordem de *Gigaflops sustained* para a resolução do problema aplicando o algoritmo de busca completa (*block matching*) a plena resolução. Observe também que a qualidade do mapa denso de disparidade (MDD) depende substancialmente desta primeira estimativa do MDD porque os mapas obtidos a maiores resoluções dependem desta primeira estimativa. Como o algoritmo de busca completa é aplicado aos ícones de baixa resolução essa é a razão pela qual somente a região de mais baixa frequência é gerada e utilizada dentre as 4 regiões geradas pelo algoritmo *Quad-tree*.

RESULTADOS E DISCUSSÃO

Para propósito de verificação da metodologia empregada foram realizados experimentos com as mesmas imagens estereoscópicas de 720 por 576 e 25 frames/segundo (Resolução utilizada no sistema Europeu de TV, *standard CCIR Rec. 601*) relatadas no algoritmo de *Chiari, Ciciani, Romero e Rossi* [2]. No total foram processadas 5 sequências estereoscópicas de cenas reais e para ilustrar o algoritmo são mostrados, neste trabalho, os resultados de exemplos de 3 classes de imagens, incluindo valores numéricos do processamento dos primeiros pares estereoscópicos: frames 0 a 9 das sequências: Trem, Carrossel e Rinoceronte.

Na Figura 1 se ilustra a imagem com poucos objetos a diferentes distâncias em profundidade (Trem). Na Figura 9 se ilustra a imagem com muitos objetos a diferentes distâncias em profundidades (Carrossel), consideradas para efeito deste trabalho como um par de imagens complexas. Perceba o número 20 e



Figura 9. Par estéreo das imagens do Carrossel: (a) Imagem esquerda, (b) Imagem direita.

a senhora indicadas pelas setas vermelhas que são objetos de interesse a serem discutidos na análise das imagens reconstruídas na seção de Resultados e Discussões. Na Figura 13 se ilustrará a imagem simples com apenas um objeto em cena (Rinoceronte).

Foram realizadas reconstruções utilizando o mapa gerado, e sobre estas reconstruções, calculada a relação sinal/ruído S/N (em dB) entre a imagem reconstruída e a original para verificar a qualidade da imagem reconstruída, o que pode ser observado por meio da análise visual dos resultados. Esta análise visual é necessária porque, nem sempre uma maior relação sinal/ruído se reflete em melhor qualidade visual.

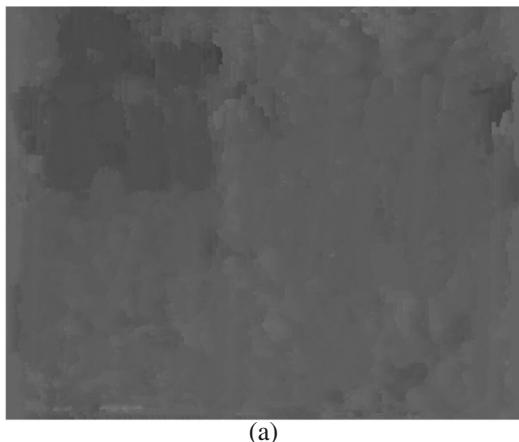
A Figura 10 apresenta os resultados obtidos utilizando a metodologia proposta com álgebra de *Clifford*, onde a Figura 10(a) é o mapa denso de disparidade do par estereoscópico presente na Figura 9, e a Figura 10(b) são as reconstruções obtidas utilizando o mapa denso de disparidade e

a imagem esquerda (Figura 9(a)), como referência. Na Tabela 6, são apresentados alguns dos valores da relação S/N obtidos utilizando luminância Y e a crominância YUV no domínio do *Clifford* para a sequência de pares estereoscópicos (0 a 9) do conjunto do Carrossel. A Figura 11 apresenta os resultados obtidos com o algoritmo proposto utilizando apenas a luminância do par estereoscópico, onde a Figura 11(a) é o mapa denso de disparidade do par estereoscópico da Figura 9, e a Figura 11(b) é a reconstrução obtida utilizando o mapa denso de disparidade e a Figura 9(a).

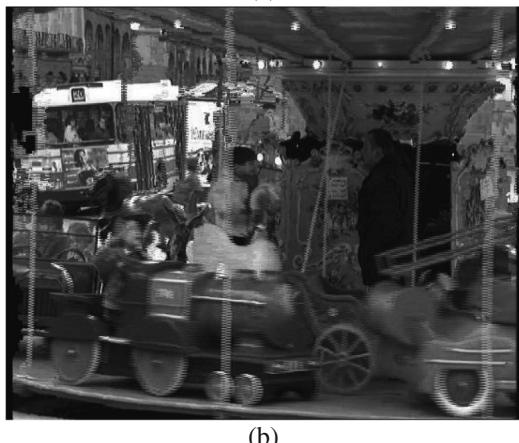
Em uma análise visual dos resultados obtidos nas imagens do Carrossel, pode se encontrar em regiões da imagem reconstruída (Figura 10(b)) algumas pessoas que estão presentes dentro do ônibus e o número que

identifica a rota do ônibus 20 que são visíveis em *Clifford* e não são visíveis na imagem reconstruída utilizando apenas luminância (Figura 11(b)), como mostrado pelas setas vermelhas. A reconstrução nas áreas oclusas, obviamente, gera resultados errados uma vez que não se pode resolver o problema da correspondência, pois o *pixel* esta presente apenas em uma das imagens direita ou esquerda. Nestes casos o método aproxima o *pixel* em questão com o *pixel* que obtiver o menor MAD.

Isto é observado na senhora que aparece do lado esquerdo na imagem direita Figura9(b), que não aparece na imagem reconstruída. Algumas das sequências de imagens usadas apresentam variação na iluminação entre o par estéreo. Isto implica que o *matching* não seja possível em algumas regiões.

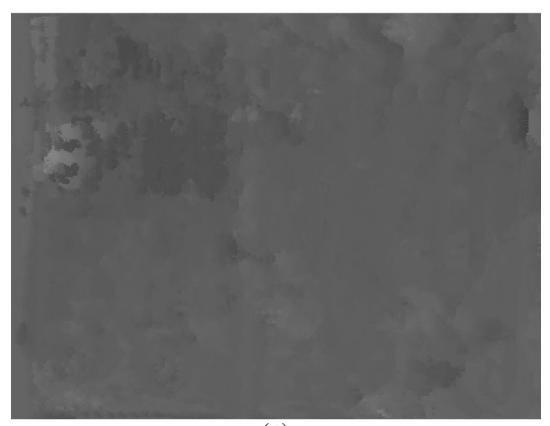


(a)

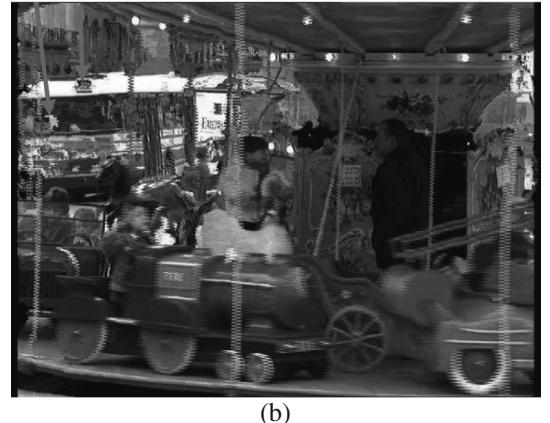


(b)

Figura 10. (a) Mapa denso da disparidade gerado utilizando álgebra de Clifford, (b) Imagem direita reconstruída utilizando o mapa denso da disparidade e a imagem esquerda (Figura 9(b)) como referência.



(a)



(b)

Figura 11. (a) Mapa denso da disparidade gerado utilizando a luminância, (b) Imagem direita reconstruída utilizando o mapa de disparidade e a imagem esquerda (Figura 9(b)) como referência.

Tabela 6. Resultados do S/N (dB) obtidos das sequências do Carrossel (0 a 9 primeiros pares estereoscópicos).

Carrossel	Y	YUV
0	27,9873	27,9003
1	28,0163	27,8934
2	27,9567	27,8980
3	27,9550	27,8514
4	27,9881	27,8976
5	27,9879	27,9211
6	28,0283	27,9608
7	28,1246	28,0789
8	28,1647	28,0746
9	28,1283	28,0474

Assim pode-se observar, nos mapas de disparidade presentes na Figura 12, que a diferença de iluminação causou o erro no *matching* onde é possível ver uma região mais escura na frente do primeiro trenzinho (parte inferior à direita), o que não acontece no mapa inferior onde a região está mais clara. Uma abordagem invariante à iluminação é tratada em [20].

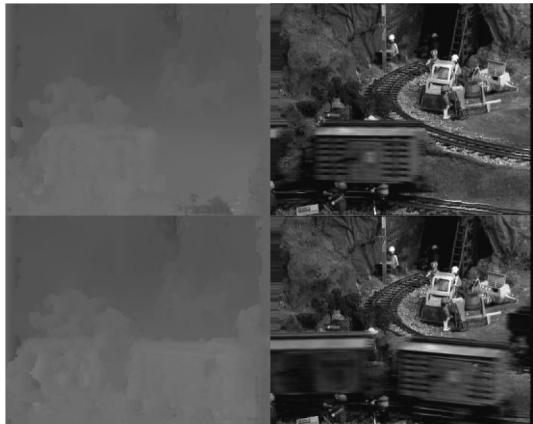


Figura 12. Resultados utilizando o YUV de diferentes cenas do *Trem*, (a) Mapa denso da Disparidade, (b) Reconstrução obtida.

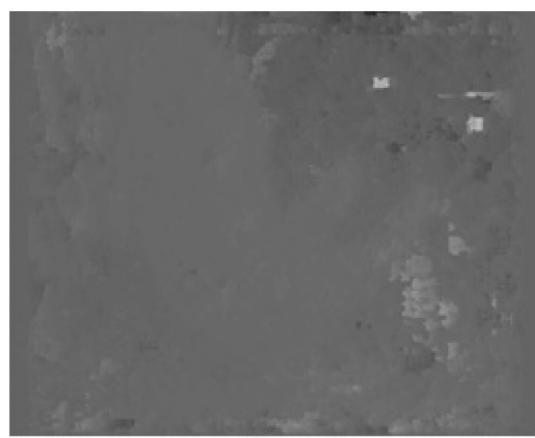
Na Tabela 7 são apresentados alguns dos valores da relação S/N obtidos utilizando luminância Y e a crominância YUV no domínio do *Clifford* para a sequência de pares estereoscópicos (0 a 9) do conjunto do Trem.

Algumas das imagens utilizadas apresentam estruturas mais simples com apenas um objeto em cena, onde os resultados utilizando *Clifford* são semelhantes

Tabela 7. Resultados do S/N (dB) obtido com as imagens do Trem (primeiros 0 a 9 pares estereoscópicos).

Trem	Y	YUV
0	30,0362	29,8461
1	29,9589	29,7863
2	29,8870	29,7754
3	29,8667	29,7946
4	29,8283	29,7518
5	29,6283	29,7227
6	29,6052	29,7156
7	29,5867	29,7100
8	29,4674	29,6443
9	29,5758	29,6956

aos obtidos utilizando a luminância, como pode ser visto na Tabela 8 e visualmente na Figura 13.



(a)



(b)

Figura 13. Resultados utilizando o Quad-tree em *Clifford* na imagem do Rinoceronte, (a) Mapa denso da disparidade, (b) Imagem reconstruída.

Tabela 8.Resultados do S/N (dB) das imagens do Rinoceronte (primeiros 0 a 9 pares estereoscópicos).

Rinoceronte	Y	YUV
0	29,6636	29,6854
1	29,6462	29,6504
2	29,6830	29,6865
3	29,6977	29,7114
4	29,7070	29,7253
5	29,7330	29,7345
6	29,7412	29,7436
7	29,7523	29,7607
8	29,7142	29,7250
9	29,7119	29,7204

Existem 4 aspectos importantes deste processamento no espaço de *Clifford* para a solução do problema da correspondência: 1) o *mapping* da crominância no domínio de *Clifford*; 2) o mapping do filtro no domínio de *Clifford*; 3) a validação dos resultados obtidos; e 4) o sistema computacional adequado para esta aplicação.

Em relação ao aspecto 1, os *pixels* da imagem utilizada (onde o YUV é mapeado em e_1, e_2, e_3) podem ser codificados de diversas formas, por exemplo, pode-se mapeá-los no escalar, vetor, bivector, etc.. Em relação ao aspecto 2, a codificação do filtro somente no escalar simplifica a computação, e note que o espaço do bi-vetor (2-vector) e o espaço do tri-vetor (3-vector) não são realmente utilizados. Adicionalmente, a codificação do filtro pode ser feita no escalar e replicado sobre os vetores, obtendo-se bastante redundância em cada dimensão do número de *Clifford*. Isto pode ajudar quando as imagens originais tenham baixa relação sinal/ruído. Fundamentalmente, cada dimensão dos vetores do domínio de *Clifford*, quando efetuada a convolução, possui uma versão filtrada em cada valor de luminância Y, crominância U ou V. Se o filtro fosse definido não somente no escalar, mas em todas as dimensões do domínio de *Clifford* teríamos a versão filtrada como antes e, adicionalmente, versões misturadas de luminância e crominâncias YU, YV, UV dependendo da forma em que se codifique o filtro.

Em relação ao aspecto 3, note que o objetivo deste trabalho demonstra a contribuição do processamento no domínio de *Clifford* em cores em relação ao

domínio real sobre a luminância para resolver o problema da correspondência.

Em relação ao aspecto 4, os resultados foram obtidos utilizando-se três programas implementados na linguagem *Java*. No primeiro é calculado e criado o mapa denso de disparidade utilizando a biblioteca *Geometric AlgeBra Learning Environment (GABLE)* [21], que possibilita o cômputo dos principais operadores das álgebras geométricas de *Clifford* de assinatura $C_{p,q}$ para o caso específico das álgebras onde $p + q = 3$. No segundo programa é realizada a reconstrução da imagem direta utilizando: 1) o mapa denso de disparidade e 2) a imagem esquerda. No terceiro programa é calculada a relação sinal/ruído na imagem reconstruída. Em relação à complexidade computacional dos algoritmos não são comparados os resultados numéricos porque o algoritmo de Chiari, Ciciani, Romero e Rossi [2] foi implementado em linguagem C e a extensão em *Clifford* na linguagem *Java* (por facilidade de utilização de bibliotecas de *Clifford*).

CONCLUSÕES E TRABALHOS FUTUROS

Foi realizada a comparação do processamento de imagens estéreo entre o algoritmo apresentado por Chiari, Ciciani, Romero e Rossi [2], que é uma extensão do *Quad-tree* para processar imagens no domínio real utilizando a luminância, com o algoritmo, aqui proposto. A extensão do *Quad-tree* aqui apresentada para estimativa do mapa denso de disparidade utilizando imagens coloridas no domínio de *Clifford* mostrou-se eficiente para imagens complexas de cenas reais. Por meio das reconstruções obtidas nesses casos é possível visualizar a melhora nos resultados, o que indica que o mapa gerado tem uma melhor representação de profundidade da cena. Em imagens de menor complexidade verifica-se que tanto utilizando *Clifford* quanto a luminância os resultados foram semelhantes. Isto mostra que a informação de cor nesses casos é pouco relevante. Aplicando *Clifford* em regiões da imagem que apresentavam diferença de iluminação entre o par estereoscópico, foram identificados valores de disparidade errôneos, indicando que a aplicação de *Clifford* ainda não resolve as variações de iluminação. O domínio de *Clifford* apresenta a vantagem adicional que um único número de *Clifford* representa as três dimensões do espaço de cores escolhido facilitando e melhorando, como

foi demonstrado experimentalmente, a estimação do mapa denso de disparidade. A alternativa para o processamento em cores seria processar cada componente por separado e isto justifica a escolha do domínio de *Clifford*. Os experimentos mostram que a ideia apresentada neste trabalho é um bom compromisso entre simplicidade (filtro codificado unicamente no escalar de *Clifford*) e os resultados obtidos.

Como trabalhos futuros, serão realizados experimentos utilizando outros modelos de cor como o *RGB*, outras formas de mapeamento das componentes de cor e do filtro de suavização no domínio de *Clifford*, a implementação em *hardware* do algoritmo, e comparar os resultados obtidos aqui em relação a outros métodos de solução do problema da correspondência.

AGRADECIMENTOS

Os autores agradecem pelo suporte financeiro ao CTEI-MS. Milton Romero foi bolsista Capes -Brasília, André Pasquali foi bolsista CNPq.

REFERÊNCIAS

- [1] D. Scharstein and R. Szeliski. "A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms". International Journal of Computer Vision. Vol. 47 Issue 1-3, pp. 7-42. 2002.
- [2] A. Chiari, B. Ciciani, M. Romero and R. Rossi. "Depth Controlled 3-DTV Image Coding". Proc. of the SPIE IS&T'98. San Jose (CA), USA. January, 1998.
- [3] M.Z. Brown, D. Burschka and G.D. Hager. "Advances in Computational Stereo". IEEE Transactions on Pattern Analysis and Machine Intelligence. Vol. 25, Issue 8, pp. 993-1008. August, 2003.
- [4] J. Banks and P. Corke. "Quantitative Evaluation of Matching Methods and Validity Measures for Stereo Vision". The International Journal of Robotics Research. Vol. 20, Issue 7, pp. 512-532. July, 2001.
- [5] U.R. Dhond and J.K. Aggarwal. "Structure from Stereo a Review". Systems, Man and Cybernetics. Vol. 19, Issue 6, pp. 1489-1510. December, 1989.
- [6] H.X. Qin, L.J. Jun and L.X. Yan. "Multi-Scale Stereo Analysis Based on Local Multi-Model Monogenic Image Feature Descriptors". International Conference on Materials Science and Information Technology (MSIT 2011). Singapore. Sep. 16-18, 2011. Singapore Inst Elect Source: Materials Science and Information Technology, Pts 1-8 Book, Series: Advanced Materials Research. Vol. 433-440, pp. 8632-859. 2012.
- [7] J. Li, H. Zhao, Ch. Shi and X. Zhou. "A Multi-Model Stereo Similarity Function Based on Monogenic Signal Analysis in Poisson Scale Space". Mathematical Problems in Engineering. Vol. 2011, Article ID 202653, p. 14. 2011. DOI: 10.1155/2011/202653.
- [8] J. Li, H. Zhao, K. Jiang, X. Zhou and X. Tong. "Multiscale Stereo Analysis Based on Local-color-phase Congruency in the Color Monogenic Signal Framework". Optics Letters. Vol. 35, Issue 13, pp. 2272-2274. 2010.
- [9] L. Yao, D. Li, J. Zhang, L. Wang and M. Zhang. "Accurate Real-time Stereo Correspondence Using Intra- and Inter-Scanline Optimization". Publisher: SP Zhejiang University Press. Journal of Zhejiang University Science C Computers&Electronics. Vol. 13, Issue 6, pp. 472-482. 2012.
- [10] M. Lang, A. Hornung, O. Wang, S. Poulakos, A. Smolic and M. Gross. "Nonlinear Disparity Mapping for Stereoscopic 3D". ACM Trans. Graph. Vol. 29, Issue 3, p. 10. 2010.
- [11] Q. Zhang and K.N. Ngan. "Dense Stereo Matching from Separated Views of Wide-Baseline Images". Editor(s): J. Blanc-Talon, D. Bone, W. Philips, D. Popescu and P. Scheunders. Conference: Advanced Concepts for Intelligent Vision Systems, Publisher: Springer Berlin Heidelberg, Booktitle: Advanced Concepts for Intelligent Vision System. Vol. 6474. Series: Lecture Notes in Computer Science, pp. 255-266. 2010. ISBN: 978-3-642-17687-6.
- [12] L. Dorst, S. Mann and G. Algebra. "A Computational Framework for Geometrical Applications (Part 1)". IEEE Computer Graphics and Applications. Vol. 22, Issue 4, pp. 58-67. July/August, 2002.

- [13] D. Hestenes. “Oersted Medal Lecture 2002: Reforming the Mathematical Language of Physics”. *American Journal of Physics*. Vol. 71, Issue 2, pp. 104-121. July, 2003.
- [14] D.N. Bhat and S.K. Nayar. “Ordinal Measures for Image Correspondence”. *Pattern Analysis and Machine Intelligence*. Vol. 20, Issue 4, pp. 415-423. April, 1998.
- [15] K. Mühlmann, D. Maier, J. Hesser and R. Männer. “Calculating Dense Disparity Maps from Color Stereo Images, an Efficient Implementation”. *International Journal of Computer Vision*. Vol. 47, Issue 1-3, pp. 79-88. April-June, 2001.
- [16] T. Bülow, M. Felsberg and G. Sommer. “Non-commutative Hypercomplex Fourier Transforms of Multidimensional Signals, in: Geometric Computing with Clifford Algebras: Theoretical Foundations and Applications in Computer Vision and Robotics”. Springer-Verlag. Ed. 1, pp. 187-207. London, UK. 2001. ISBN:3-540-41198-4.
- [17] J. Ebling and G. Scheuermann. “Clifford Convolution And Pattern Matching On Vector Fields”. *VIS ‘03: Proceedings of the 14th IEEE Visualization 2003 (VIS’03)*. Seattle. Washington, USA. October, 2003.
- [18] B. Mishra, P. Wilson and B. Al-Hashimi. “Advancement in Color Image Processing using Geometric Algebra”. *16th European Signal Processing Conference (EUSIPCO)*. Lausanne, Switzerland. August, 2008.
- [19] M.Z. Brown, D. Burschka and G.D. Hager. “Color Edge Detection Hardware Based on Geometric Algebra”. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. Vol. 15, Issue 6, pp. 1454-1463. April, 2006.
- [20] L. Nalpantidis and A. Gasteratos. “Stereo Vision for Robotic Applications in the Presence of Non-ideal Lighting Conditions”. *Image and Vision Computing*. Vol. 28, Issue 6, pp. 940-951. June, 2010. ISSN 0262-8856. DOI: 10.1016/j.imavis.2009.11.011.
- [21] S. Mann, L. Dorst and T. Bouma. “The Making of GABLE: a Geometric AlgeBra Learning Environment in Matlab”. In: *Geometric Algebra with Applications in Science and Engineering*. Birkhauser, Ed. 1. Boston, MA, USA. Vol. 1, pp. 491-511. 2001. ISBN: 0-8176-4199-8.