

Correspondência automática de imagens coloridas utilizando a técnica de correlação com ponderação adaptativa.

Paulo Henrique Amorim da Silva¹
Antonio Maria Garcia Tommaselli²
Maurício Galo²

¹ Bolsista IC/Fapesp

² Universidade Estadual Paulista – UNESP - Campus de Presidente Prudente.

Rua Roberto Simonsen, 305, Jd. Das Rosas, CEP 19060-900, CP 468

¹amorim.cart@gmail.com, ²{tomaseli.galo}@fct.unesp.br.

ABSTRACT

Area-based matching is one of the most important photogrammetric process in Digital Photogrammetry, because it allows the automation of several stages of the photogrammetric pipeline. In most of the commercial softwares nowadays available, the algorithms of image correlation uses only the intensity information (gray levels), despising the color information, that could be useful, if used in an appropriate way, increasing the efficiency of the current correspondence techniques in Digital Photogrammetry. The aim of this work is to present a technique that use the IHS color model in the correlation process, were each color channel and three correlation matrixes are generated. The results are presented and discussed, showing that using this technique increases the robustness of the correlation process.

Palavras Chaves: Fotogrametria digital, correspondência de imagens, Visão computacional, digital Photogrammetry, image matching, Computer Vision.

1. Introdução

A Fotogrametria digital tem suas raízes por volta de 1950, com trabalhos sobre correlação de imagens, mas foi somente na década de 1980, devido aos avanços eletrônicos e computacionais, que os processos fotogramétricos tornaram-se mais ágeis, iniciando sua evolução para os meios digitais (SCHENK, 1999). A determinação da correspondência entre pontos homólogos é o procedimento fotogramétrico que tem mais relevância na fotogrametria digital, pois permite a automação das etapas de orientação relativa das imagens, fototriangulação e a determinação da posição 3D dos pontos no espaço objeto (geração de Modelos Digitais de Terreno).

A correlação de imagens digitais mostra-se como uma etapa relevante, uma vez que o sucesso desta etapa é primordial para se obter resultados confiáveis nos processos fotogramétricos posteriores, como a fototriangulação e a geração de Modelos Digitais de Terreno.

Na grande maioria dos softwares comerciais em uso, os algoritmos de correlação automática operam basicamente com a informação de intensidade (ou tons de cinza), desprezando as informações relativas à cor, que podem ser úteis, se utilizadas de modo adequado no aumento do desempenho das técnicas de correspondência.

Diante disso, nota-se a importância de se desenvolver e estudar métodos de correspondência automática de imagens que explorem a informação das cores, na tentativa de obter melhores resultados nos processos fotogramétricos.

Paradoxalmente, este tema é pouco abordado na literatura fotogramétrica e alguns poucos trabalhos utilizam as informações extraídas dos canais de cores, principalmente no contexto de correspondência por feições (BROCKELBANK e YANG, 1989; HUNG et al., 1998; PARK e ZINMMERMANN, 2000; CHAMBON e CROUZI, 2004, 2005).

2. Fundamentação teórica

Segundo Andrade (1998), a detecção de pontos homólogos em imagens estereoscópicas consiste em fazer com que uma matriz de amostra de uma das imagens percorra toda a região de uma janela (matriz de busca) na imagem homóloga, identificando a região mais propícia à região de correspondência.

Uma das funções mais utilizadas para a medida de similaridade é o coeficiente de correlação, que retorna um número no intervalo -1 e +1. O valor 1 corresponde à medida de similaridade máxima, o valor 0 indica que não há correspondência entre as janelas e o valor -1 indica correlação inversa. A expressão que permite o cálculo do coeficiente de correlação é apresentada em (1).

$$\rho = \frac{\sigma_{RB}}{\sigma_R \sigma_B} = \frac{\sum_{i=1}^n \sum_{j=1}^m ((g_{R(i,j)} - \bar{g}_R)(g_{B(i,j)} - \bar{g}_B))}{\sqrt{\sum_{i=0}^n \sum_{j=0}^m (g_{R(i,j)} - \bar{g}_R)^2 \cdot \sum_{i=0}^n \sum_{j=0}^m (g_{B(i,j)} - \bar{g}_B)^2}} \quad (1)$$

sendo:

σ_{RB} - a covariância entre os tons de cinza das janelas de referência e de busca;

σ_R - o desvio-padrão dos tons de cinza na janela de referência;

σ_B - o desvio-padrão dos tons de cinza na janela de busca; n, m - são as dimensões da janela de referência;

$g_{R(i,j)}$ - o nível de cinza na posição (i,j) na janela de referência;

$g_{B(i,j)}$ - o nível de cinza da posição (i, j) na janela de busca;

\bar{g}_R - a média dos níveis de cinza da janela de referência; e

\bar{g}_B - a média dos níveis de cinza da janela de busca.

Segundo Lillesand e Kiefer (2000), as cores amarelo, magenta e ciano, formam o sistema YMC que é utilizado em dispositivos como impressoras. É possível representar todas as cores visíveis através da combinação das três componentes primárias, R, G e B formando o sistema de cor RGB, utilizado para compor as cores em dispositivos como monitores e câmaras digitais.

Outra forma de decompor as cores é pelo sistema HSI. Neste modelo as cores são separadas nas componentes: Matiz (*Hue*), Saturação (*Saturation*) e Intensidade (*Intensity*). Segundo Centeno (2003), este modelo é uma representação muito utilizada, pois a componente intensidade (I) pode ser desvinculada da informação de cor de uma imagem (matiz e saturação); e suas componentes de matiz (H) e saturação (S) estão intimamente relacionadas com o processo pelos quais os seres humanos percebem a cor. Deste modo, cada pixel será representado em três bandas (H, S, I), a partir da transformação RGB-HSI, e cada banda será tratada separadamente. Neste modelo, as características usadas para distinguir uma

cor de outra são suas componentes de matiz ou tonalidade, saturação e intensidade, conforme ilustrado na Figura 1.

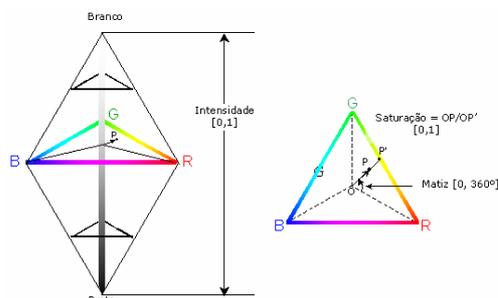


Figura 1 – O modelo HSI em relação ao modelo RGB.
(Fonte: Adaptado de Gonzalez e Woods, 2000).

2.1. Estimativa da qualidade da correlação de áreas

A qualidade da correlação pode ser estimada antes mesmo de realizá-la, calculando-se a Matriz Variância-Covariância das translações entre as janelas de referência e pesquisa. Em síntese, este processo pode ser realizado considerando as seguintes etapas (COSTA et al., 2005; HARALICK e SHAPIRO, 1993): primeiramente estima-se a variância do ruído em função da variância da janela de referência e do máximo coeficiente de correlação; a seguir, calcula-se a matriz de equações normais em função dos gradientes nas linhas e colunas da janela de referência e finalmente calcula-se a precisão dos parâmetros de translação entre as imagens em função da variância do ruído e da matriz de equações normais.

Deste modo, pode-se observar que a MVC das translações permite que se tenha uma estimativa da qualidade da correlação, baseado apenas nas informações de intensidade da janela de referência. Por exemplo, uma borda horizontal perfeita não permite a determinação da translação em x, ao passo que um canto permite boa qualidade, tanto ao longo das linhas quanto colunas. Assim, o traço da MVC das translações, obtido pela soma dos elementos de sua diagonal principal, pode variar no intervalo 0 até $+\infty$, sendo que uma imagem boa para a correlação apresenta o traço da MVC das translações muito próximo de 0 (COSTA et al., 2005).

Costa et al. (2005) verificaram empiricamente que valores maiores que 0,09 podem ser rejeitados, pois se mostram deficientes para a correlação, como no caso de regiões homogêneas ou bordas horizontais.

3. Metodologia

As imagens utilizadas nos testes correspondem a janelas recortadas em um estereopar de imagens aéreas digitais coloridas contendo uma parte da cidade de Assis - SP, capturadas por uma câmara digital Hasselblad Imacon H1D (22 Mpxels).

Utilizou-se nos experimentos janelas de referência de dimensões 9x9 e 7x7 pixels, pois sabe-se que ocorrem distorções na geometria das imagens que prejudicam o processo de correspondência para janelas de referência de grandes dimensões. Já para as janelas de busca, foram utilizadas janelas de lados variando de 69 a 97 pixels. A Figura 2 ilustra as imagens das janelas de busca e referência selecionadas. O ponto correspondente foi destacado com um círculo sobre a janela de busca.

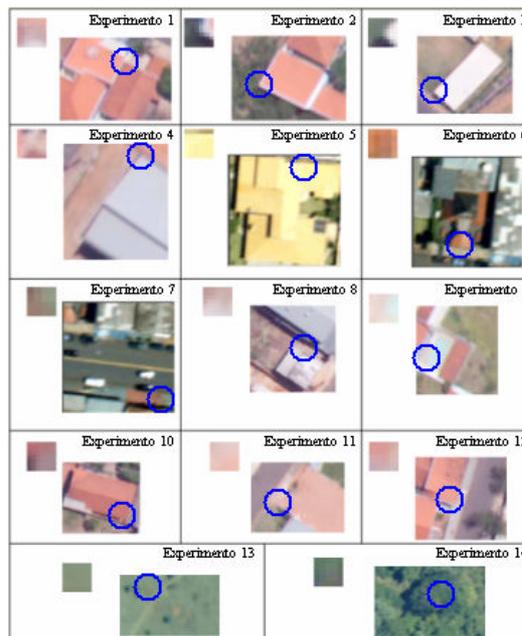


Figura 2 - Janelas de referência e de busca utilizadas nos experimentos.

Neste trabalho foi desenvolvida uma estratégia para o estabelecimento da correspondência por correlação usando imagens aéreas digitais coloridas, visando minimizar eventuais problemas de falsas correlações que ocorrem nas imagens em tons de cinza. Para isto, foi desenvolvido um programa em linguagem C++, utilizando o compilador *Borland C++Builder*, cuja finalidade é fazer a leitura dos canais primários de cor de cada pixel das imagens aéreas digitais, no formato ppm (*portable pixelmap*), que são posteriormente processadas com os algoritmos desenvolvidos.

Com o objetivo de estudar a influência das cores no processo de correspondência, o programa realiza tanto a correlação em tons de cinza, como também nas imagens decompostas nos modelos RGB e HSI.

3.1. Correspondência usando o modelo RGB

Para realizar a correlação utilizando a informação das cores, é necessário, primeiramente, decompor as imagens de busca e de referência nos três canais (R, G, B), para depois aplicar o processo de correlação, gerando-se três matrizes (ou imagens) de correlação (uma para cada canal). Finalmente, calcula-se uma imagem de correlação média, através da média aritmética ou da média ponderada, na qual são atribuídos pesos proporcionais à qualidade esperada para a correlação.

Primeiramente foram realizados testes calculando-se o coeficiente médio de correlação, usando a média aritmética das imagens de correlação dos três canais para o modelo RGB e para o modelo HSI. Nota-se que utilizando este método os resultados obtidos foram melhores que utilizando somente a intensidade ou tons de cinza. No entanto, alguns experimentos ainda apresentaram falsas correlações, o que motivou o desenvolvimento de um método de ponderação adaptativa.

A idéia central é que um canal adequado para a correlação tenha traço pequeno, ao qual se associaria um peso alto e, inversamente, a uma área inadequada se associaria um peso próximo de zero. Na prática, entretanto, observa-se que a variação do traço de um canal bom para correlação e de um canal ruim é muito pequena. Assim, deve-se usar uma função de

similaridade que permita esta discriminação, associando um valor de similaridade alto quando os traços são pequenos.

Observando este comportamento, pode-se notar que ele tem o comportamento da função usada por Faugeras (1993) e Galo (2003) para o estabelecimento da correspondência, usando o método de rotulação por relaxação. Esta função pode ser escrita por:

$$P_{\{R,G,B\}} = \frac{1}{1 + \alpha \cdot tr_{\{R,G,B\}}} \quad (2)$$

sendo α uma constante relacionada com a influência da medida de distância, ou traço ($tr_{\{R,G,B\}}$), neste caso.

Uma outra função que pode ser considerada é uma curva gaussiana, cuja função é:

$$P_{\{R,G,B\}} = e^{-0.5 \left(\frac{tr_{\{R,G,B\}}}{\sigma} \right)^2} \quad (3)$$

sendo σ uma constante relacionada com a forma da curva.

Após realizar vários testes, foi determinado um valor empírico para σ e α . Para se chegar a tais valores foram realizados testes com 7 das 15 imagens disponíveis para testes, variando-se o valor de σ e α até atingir o melhor resultado. Aplicou-se, então, o valor obtido de $\sigma=0,005$ e $\alpha=40$ para os experimentos restantes, nos quais se pôde observar o bom desempenho destas funções, as quais apresentaram os mesmos resultados nos experimentos realizados.

3.2. Correspondência usando o modelo HSI

Como as características das cores variam de imagem para imagem, desenvolveu-se uma metodologia de atribuição automática de pesos para o modelo HSI, assim como no modelo RGB ponderado. Este método baseia-se no canal S da imagem, que varia no intervalo de 0 a 1. Logo, quanto maior for a saturação média, mais nítidas serão as cores da imagem. Assim, maior deverá ser o peso atribuído aos canais H e S e menor ao canal I.

Apesar dos bons resultados obtidos no modelo RGB, o traço da MVC não foi usado como argumento para o cálculo do peso no modelo HSI, devido às suas características, como foi visto na seção 2, e ao fato da componente H (matiz) ser um ângulo. Por esta razão, aplicou-se a determinação automática dos pesos dos canais H, S e I em função apenas da saturação.

A partir desta hipótese foi formulada empiricamente uma função matemática que descreve o comportamento do peso em função da saturação média, função esta que foi posteriormente testada. A função utilizada para descrever tal comportamento pode ser dada por:

$$P_{\{H,S\}} = (S_{m\u00e9dia})^{1/n} \quad \text{e} \quad P_{\{I\}} = 1 - (S_{m\u00e9dia})^{1/n} \quad (4)$$

Nota-se que quando a saturação média tende a 1, o peso das imagens de correlação dos canais H e S tendem a 1, enquanto o peso da imagem de correlação de I tende a 0. Quando a saturação média tende a 0, o peso das imagens de correlação H e S também tendem a zero, enquanto o peso de I tende a 1.

Após realizar vários testes, foi determinado um valor empírico para n . Para se chegar ao valor de n , foram realizados testes com 7 das 15 imagens disponíveis para testes, variando-se

o valor de n até atingir o melhor resultado. Aplicou-se, então, o valor obtido de $n=2$ para os experimentos restantes, nos quais se pôde observar o bom desempenho desta função.

Para se calcular a MVC das translações de uma imagem é necessário estimar uma variância do ruído a priori. A partir dos experimentos realizados com as imagens aéreas da câmara Hasselblad, Costa et al. (2005) recomenda utilizar uma variância à priori de 0,9. Como as imagens usadas nos experimentos deste trabalho são as mesmas usadas por Costa et al. (2005), foi adotado este mesmo valor.

Para efeito de comparação, os mesmos dados foram usados para realizar a correlação com as imagens transformadas para tons de cinza, como apresentado na seqüência.

As figuras a seguir apresentam dois gráficos. O primeiro ilustra o erro na correspondência em cada experimento, utilizando correlação nas imagens em tons de cinza e o método de média aritmética nos coeficientes de correlação nos modelos de cor RGB e HSI. O segundo gráfico apresenta uma comparação entre a correlação utilizando imagens em tons de cinza e a correlação utilizando a técnica de ponderação adaptativa para os modelos RGB e HSI. Nota-se uma diminuição expressiva dos erros de correspondência em praticamente todos os experimentos, o que significa quase ausência de falsas correlações ao usar as técnicas desenvolvidas.

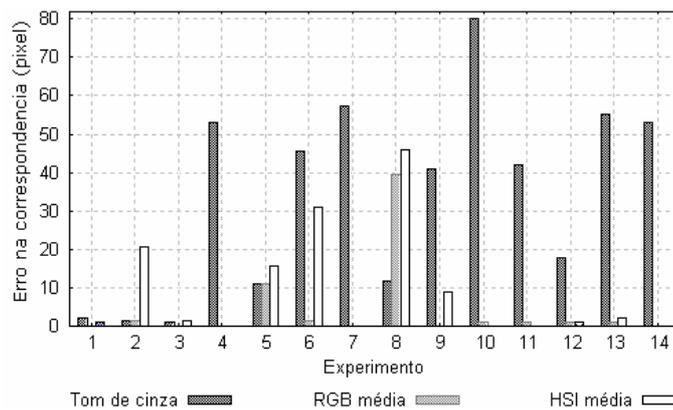


Figura 3 - Erro da correspondência utilizando tons de cinza, média aritmética nos modelos RGB e HSI.

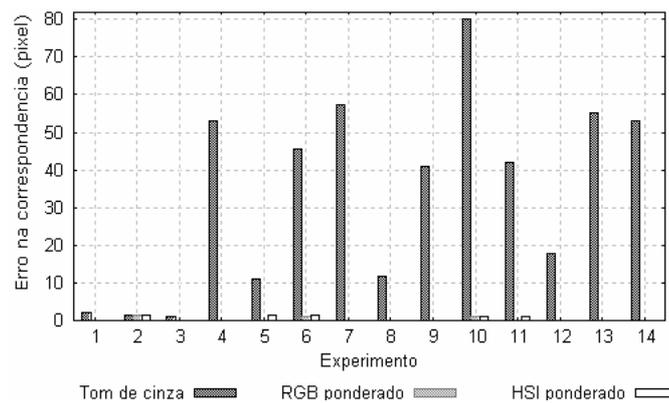


Figura 4 - Erro da correspondência usando tons de cinza, e distribuição adaptativa nos modelos RGB e HSI.

Nota-se que a correlação utilizando tons de cinza não apresentou resultados satisfatórios podendo-se observar erros que chegam até 80 pixels.

Uma vez que os valores dos erros obtidos a partir dos métodos RGB e HSI adaptativo (ou ponderado) são pequenos e não podem ser vistos na Figura 4, apresenta-se a Figura 5, na qual somente os resultados com a correlação nos canais coloridos são mostrados.

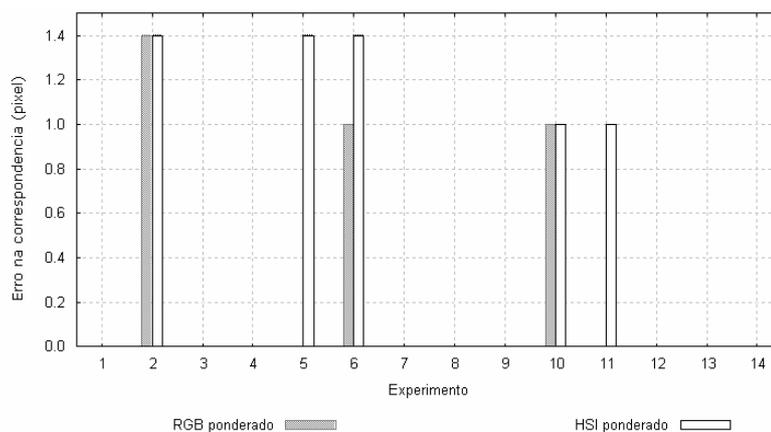


Figura 5 - Erro da correspondência utilizando os métodos RGB e HSI adaptativo.

4. CONCLUSÃO

Os resultados obtidos neste trabalho permitem concluir que é melhor usar os vários canais de cores do que somente a intensidade (ou tons de cinza) no processo de estabelecimento da correspondência de áreas a partir de imagens digitais, pois o uso da intensidade apenas, deixa de incluir informações necessárias para resolver ambigüidades que podem causar falsas correspondências.

Observa-se que na correlação utilizando tons de cinza, o erro médio obtido nos experimentos foi de 30 pixels, podendo-se notar alguns erros de grandes magnitudes em janelas onde as imagens são ruins para o estabelecimento da correspondência (homogêneas ou com ambigüidade). O uso da informação das cores dos modelos RGB e HSI, através da média aritmética dos coeficientes de correlação de cada um desses canais melhorou estes resultados de forma expressiva, no entanto apresentavam erros significativos em alguns experimentos. A utilização da média ponderada ou adaptativa dos coeficientes de correlação dos modelos de cores RGB e HSI tornou o processo mais robusto, ou seja, menos vulnerável a falsas correlações, obtendo-se um erro médio de 0,2 pixels em ambos os modelos, sendo o erro máximo, nos experimentos realizados, de 1,4 pixels. Por outro lado, os pequenos erros resultantes, da ordem de 1,4 pixels podem ser eliminados utilizando técnicas de correlação subpixel (SCHENK, 1999).

Os resultados obtidos se mostram promissores, pois nota-se a possibilidade de aumentar a qualidade e eficiência do processo de correspondência automática de imagens pelo uso da ponderação adaptativa dos coeficientes de correlação de cada componente do espaço de cores RGB. O peso utilizado é calculado em função do traço da MVC das translações da janela de referência.

Recomenda-se para trabalhos futuros estudar esta abordagem de ponderação adaptativa, procurando avaliar se a variância do ruído é a mesma para todos os canais, tornando o processo de cálculo menos dependente de informações a priori. Adicionalmente, sugere-se a realização de um número maior de experimentos, com diferentes tipos de alvos e imagens, visando uma análise estatística mais rigorosa.

Agradecimentos

Os autores desejam expressar seus agradecimentos à **Fapesp**, pelo financiamento a este projeto na forma de uma bolsa de Iniciação Científica e à Empresa Engemap Engenharia, Mapeamento e Aerolevanteamento Ltda, pela cessão das imagens usadas nos experimentos. Uma versão preliminar deste trabalho foi apresentada no XVIII Congresso de Iniciação Científica da Unesp, 2006.

Referências

- ANDRADE, J. de B. **Fotogrametria**. Curitiba: SBEE, 246 p. 1998.
- BROCKELBANK, D. C.; YANG, Y. H. An Experimental Investigation in the Use of Color in Computational Stereopsis. **IEEE Transactions on Systems, Man and Cybernetics**, v.19, n. 6, p.1365-1383, 1989.
- CENTENO, J. A. S. **Sensoriamento Remoto e Processamento Digital de Imagens**. Curitiba: Curso de Pós Graduação em Ciências Geodésicas, 2003.
- CHAMBON, S.; CROUZIL, A. Colour correlation-based matching. In: **International Journal of Robotics and Automation**, Iasted, v. 20, n. 2, p. 78-85, 2005.
- CHAMBON, S.; CROUZIL, A. Colour stereo matching using correlation measures. In: **Complex Systems Intelligence and Modern Technological Applications – CSIMTA 2004**, p. 520-525, Chesbourg, France, 2004.
- COSTA, E. R. ; TOMMASELLI, A. M. G ; GALO, M. Técnicas de pré-análise do processo de correlação visando a geração de modelo digital de terreno. **Revista Brasileira de Cartografia**, Rio de Janeiro, v. 57, n. 3, p. 252-262, 2005.
- FAUGERAS, O. **Three-Dimensional Computer Vision - A geometric viewpoint**. Cambridge, England: The MIT Press, 663 p. 1993.
- GALO, M. **Automação dos processos de correspondência e orientação relativa em visão estéreo**. 2003. 262 p. Tese (Doutorado em Engenharia Elétrica) - Faculdade de Engenharia Elétrica e de Computação - FEEC, Universidade Estadual de Campinas - UNICAMP, Campinas, 2003.
- GONZALEZ, R. C.; WOODS, R. E. **Processamento de Imagens Digitais**. São Paulo: Editora Blücher Ltda, 703 p. 2000.
- HARALICK, R.; SHAPIRO, L. G. **Computer and robot vision - vol. II**. Addison-Wesley Publishing Company, Inc., 630, p. 1993.
- HUNG, Y. P.; CHEN, C. S.; HUNG, K. C.; CHEN, Y. S.; FUH, C.S. Multipass Hierarchical Stereo Matching for Generation of Digital Terrain Models from Aerial Images. **Machine Vision and Applications**, v. 10. p. 280-291, 1998.
- SCHENK, T. **Digital Photogrammetry – Volume I**. Laurelville: TerraScience, 428 p., 1999.
- PARK, H.J.; ZINMMERMANN, P. Colour image matching for DTM generation and house extraction. In: ISPRS Congress, 2000, Amsterdam. **International Archives of Photogrammetry and Remote Sensing**, Amsterdam. v. XXXIII, p. 697-704, 2000.