

CLASSIFICAÇÃO AUTOMÁTICA DE PADRÕES PROJETADOS POR UM SISTEMA DE LUZ ESTRUTURADA

Automatic classification of targets projected with a structured light system

CHRISTIANE NOGUEIRA DE CARVALHO KOKUBUM
ANTONIO MARIA GARCIA TOMMASELLI
MÁRIO LUIZ LOPES REISS

Universidade Estadual Paulista
Faculdade de Ciências e Tecnologia
Departamento de Cartografia
Rua Roberto Simonsen, 305 CEP 19060-900 Presidente Prudente, SP, Brasil
{kokubum, tomaseli, mreiss}@prudente.unesp.br

RESUMO

Um dos problemas mais significativos em Visão Computacional e em Fotogrametria Digital é a reconstrução 3D. A reconstrução 3D por luz estruturada é uma das técnicas existentes e que apresenta vários problemas, um dos quais a identificação ou classificação dos padrões projetados. O tratamento deste problema é o objetivo deste trabalho. Para a classificação de tais padrões, usou-se o método baseado em área denominado correspondência por padrão (*template matching*). Este método consiste na classificação por correlação, que mede a similaridade entre as janelas de referência e de busca, utilizando uma função de correlação adequada. A função usada, neste trabalho, é a covariância cruzada modificada ou coeficiente de correlação, e que foi a que proporcionou melhores resultados. Foi desenvolvida uma estratégia para reamostragem adaptativa dos padrões, o que resolveu o problema de deformação dos alvos devido à inclinação da superfície. Experimentos com dados simulados e reais foram conduzidos com o objetivo de verificar a eficácia da metodologia para a identificação dos alvos projetados. Os resultados mostram que a metodologia de classificação funciona adequadamente, identificando 98% de alvos em superfícies planas e 93% em superfícies oblíquas.

ABSTRACT

One of the main problems in Computer Vision and Close Range Digital Photogrammetry is 3D reconstruction. 3D reconstruction with structured light is one of the existing techniques and which still has several problems, one of them the identification or classification of the projected targets. Approaching this problem is the goal of this paper. An area based method called template matching was used for target classification. This method performs detection of area similarity by correlation, which measures the similarity between the reference and search windows, using a suitable correlation function. In this paper the modified cross covariance function was used, which presented the best results. A strategy was developed for adaptative resampling of the patterns, which solved the problem of deformation of the targets due to object surface inclination. Experiments with simulated and real data were performed in order to assess the efficiency of the proposed methodology for target detection. The results showed that the proposed classification strategy works properly, identifying 98% of targets in plane surfaces and 93% in oblique surfaces.

1. INTRODUÇÃO

Atualmente, a reconstrução tridimensional é um dos assuntos mais abordados na literatura relacionada à Fotogrametria Digital e a Visão Computacional. Há diversos sistemas comerciais para reconstrução 3D no mercado, porém, apresentam algumas das seguintes desvantagens: alto custo; necessidade de operadores altamente qualificados e treinados; impossibilidade de resposta em tempo real em um curto período de tempo e dificuldade de reconstrução em superfícies com texturas homogêneas.

Dentre os sistemas que minimizam alguns destes problemas pode-se identificar os sistemas de luz estruturada, que utilizam um projetor de padrões e uma câmara digital.

Os principais problemas nestes sistemas estão relacionados à calibração do sistema câmara-projetor e à identificação e localização precisa dos padrões projetados. O objetivo principal deste trabalho é apresentar algumas metodologias estudadas para o problema de identificação dos alvos. Os problemas de calibração e localização não serão tratados neste trabalho. Será dada ênfase aos resultados experimentais obtidas pelas diferentes técnicas estudadas de localização. A revisão bibliográfica e descrição de técnicas de processamento das imagens também serão feitos brevemente. Serão descritas e comparadas três metodologias para a classificação dos alvos, todas usando como base o método de correspondência por padrão (*template matching*), que procura medir a similaridade entre as máscaras de referência e de busca utilizando uma função de correlação adequada. Foram definidas variantes do método, que foram testadas com dados reais.

2. REVISÃO BIBLIOGRÁFICA

A implementação de um sistema de luz estruturada requer a integração de conhecimentos de várias áreas, particularmente de Fotogrametria, Visão de Máquina e Processamento de Imagens. Existem distintas possibilidades para a solução dos problemas nas várias etapas de construção do sistema, tanto dos componentes de *hardware* quanto de *software*, que implementam os métodos de processamento, análise e reconstrução das imagens.

Nas seções seguintes serão apresentados alguns tópicos de interesse que não esgotam a literatura sobre o assunto, mas apresentam apenas algumas das possibilidades para processamento. Uma revisão abrangente sobre todos os métodos disponíveis para o processamento das imagens está além do escopo deste trabalho.

2.1 Sistemas de Reconstrução por Luz Estruturada

Segundo Newton e Mitchell (1996), a técnica de reconstrução por luz estruturada envolve o uso de um projetor e uma câmara, na qual o sistema de projeção é usado para projetar o padrão sobre o objeto na cena e uma câmara captura a imagem dos padrões que são distorcidos pela superfície de projeção. O projetor é equivalente à segunda câmara usada em Fotogrametria e qualquer ponto projetado na cena com direção e identificação conhecidas pode ser reconstruído usando os princípios de intersecção dos raios de luz da Fotogrametria analítica (relação câmara-projetor). A determinação das relações geométricas existentes entre o sensor e os feixes de raios luminosos do padrão projetado é feita por um processo de calibração. O sistema de luz estruturada captura uma superfície em particular para cada ponto de vista, sendo necessária a combinação de vários pontos de vistas para criar um modelo 3D completo do objeto. Uma das vantagens deste sistema é quando este usa uma fonte de luz branca e, conseqüentemente, não há a necessidade de cuidados especiais de segurança, ao contrário de sistemas de varredura utilizando *laser*, cuja luz possui algum nível de radiação. Além disso, alguns sistemas de luz estruturada possuem precisões equivalentes às do sistema a laser. Como vantagem adicional deste sistema há a possibilidade de se capturar todos os alvos projetados e a imagem do objeto de uma só vez, sendo desnecessário manter o objeto estático por um longo período de tempo. Também pode haver a necessidade de tomada de múltiplas posições para cobrir inteiramente um objeto, quando ele for maior que o campo de cobertura do sensor.

As vantagens deste sistema em relação a técnicas estereofotogramétricas estão relacionadas com (BATTLE et al, 1996):

- menor tempo entre a tomada de fotos e a resposta final;
- menor custo dos equipamentos;
- não necessidade de pessoas especializadas para manusear os equipamentos; e,

– menor dificuldade de medir superfícies homogêneas.

Embora os sistemas de reconstrução por luz estruturada minimizem o problema de iluminação do ambiente ao projetar luz estruturada, ainda são necessárias várias operações de realce das imagens, até as etapas de reconhecimento e medição precisa dos alvos projetados. Distintas alternativas de processamento estão disponíveis, como mostra a farta literatura na área, tendo sido exploradas neste trabalho apenas algumas destas técnicas.

2.2 Realce do Contraste

O realce de contraste linear consiste no processamento de uma imagem, de modo que o resultado seja mais apropriado para uma aplicação específica do que a imagem original. O contraste nas imagens é afetado pela iluminação não homogênea e abertura incorreta do diafragma da câmara durante a aquisição da imagem. É normalmente usado como uma etapa de pré-processamento para sistemas de reconhecimento de padrões (GONZALEZ e WOODS, 2000).

Dentre os métodos existentes para solucionar este problema encontram-se: método de expansão do histograma, método de contraste linear percentual e método de transformação local.

Segundo Gonzalez e Woods (2000), o método de expansão do histograma, ocupa todo o intervalo de cinza, não provocando nenhuma perda. Esta operação é realizada a partir de uma interpolação linear (Equação 1).

$$G_i = \left(\frac{G_i - G_{\min}}{G_{\max} - G_{\min}} \right) * maior_nível \quad (1)$$

Onde:

G_{\min} – menor nível de cinza encontrado na imagem;

G_{\max} – maior nível de cinza encontrado na imagem;

G_i – nível de cinza original de um *pixel* na imagem;

G_i – nível de cinza resultante de um *pixel* após a transformação;

maior_nível – indica o maior nível de quantização disponível. Por exemplo: se a imagem possui 8 bits, o maior nível é dado por $2^8 - 1 = 255$.

Outra forma de melhoramento do contraste é denominada de contraste linear percentual. Este processo é utilizado quando se deseja especificar os valores de G_{\min} e G_{\max} para realçar determinado intervalo de cinza. Neste caso, utiliza-se certo percentual de *pixels* em relação à média do histograma. O problema deste processo é a perda de informação em benefício do realce dos *pixels* no intervalo que se deseja (GONZALEZ e WOODS, 2000).

Os métodos mostrados anteriormente são globais no sentido de que os *pixels* são modificados através de uma função de transformação baseada na distribuição dos níveis de cinza sobre uma imagem inteira, não garantindo o realce local desejado. Uma transformação local baseada nas propriedades de intensidade da média e do desvio-padrão de *pixels* é definida pela Equação 2 (GONZALEZ e WOODS, 2000).

$$g(x, y) = A(x, y)[f(x, y) - m(x, y)] + m(x, y) \quad (2)$$

Onde:

$$A(x, y) = kM / \sigma(x, y);$$

$f(x, y)$ - níveis de cinza da imagem original;

$g(x, y)$ - níveis de cinza da imagem resultante;

$m(x, y)$ - nível de cinza médio calculado numa vizinhança centrada em (x, y) ;

$\sigma(x, y)$ - desvio-padrão calculado numa vizinhança centrada em (x, y) ;

M - média global de $f(x, y)$;

k - constante variando no intervalo $[0, 1]$.

Este método apresenta melhores resultados em relação ao método de expansão do histograma pelo fato de realizar um melhoramento local, porém, possui uma desvantagem para a automação do processo, pois a constante k a ser definida depende das características da imagem original.

2.3 Segmentação e Crescimento de Regiões

A segmentação consiste em subdividir uma imagem em regiões homogêneas usando propriedades escolhidas, como o nível de cinza, cor, textura, entre outras (SONKA et al, 1998). Os principais objetivos da segmentação são reduzir a quantidade de dados a serem processados na análise de imagens e obter informações importantes a respeito de objetos e feições nelas contidos (EL-HAKIM, 1996). Existem diferentes taxonomias para classificar os métodos existentes de segmentação. Sonka et al (1998) dividem os métodos de segmentação em três categorias: baseados em bordas, regiões ou no conhecimento global sobre a imagem. Ballard e Brown (1982) consideram dois tipos de segmentos principais, bordas e regiões, que podem ser separados por diferentes critérios. Jain (1989) separa as técnicas de segmentação em seis grupos distintos: limiarização, rotulação de componentes, fronteiras, regiões, correspondência por padrões e textura.

Os métodos de segmentação baseados em regiões procuram dividir a imagem em segmentos que tem características similares (JAIN, 1989). Um dos métodos mais conhecidos para este tipo de segmentação é o crescimento de regiões. Este método consiste em agrupar *pixels* vizinhos que representam regiões homogêneas

segundo algum critério de similaridade. Um dos critérios que pode ser usado é a comparação da média e do desvio-padrão dos níveis de cinza de duas regiões, verificando-se se estes valores são estatisticamente compatíveis.

Outros métodos de segmentação por regiões podem ser citados, como o de subdivisão e fusão e a segmentação *watershed*, baseada nos conceitos da morfologia matemática.

2.4 Limiarização

A operação de limiarização refere-se à definição de um valor ou de uma função, que permita separar elementos de interesse na imagem do fundo. Frequentemente a imagem original é convertida em imagem binária pelo processo de limiarização, atribuindo-se o valor 0 aos pixels com tons de cinza menores que o limiar e 1 aos maiores. Este método pode ser considerado como um dos mais simples para a segmentação de imagens.

A limiarização pode ser feita usando técnicas globais ou locais. Nas técnicas globais determina-se um único limiar para toda a imagem, o que traz problemas causados pela reflectância, sombras ou iluminação inadequada. Em imagens digitalizadas com *scanners*, particularmente de documentos, as técnicas globais podem ser apropriadas, porque o histograma da imagem é bimodal e similar em todas as regiões.

Em imagens coletadas com câmaras digitais de cenas 3D, entretanto, ocorre o efeito de escurecimento nas regiões periféricas da imagem. Devido à lei de iluminância, a luz que atinge os cantos da imagem é atenuada por um fator de $\cos^4\theta$, onde θ é o ângulo entre o eixo ótico da câmara e o ponto em análise. Com isto a área central da imagem é mais iluminada que a periferia, gerando um efeito análogo ao de vinhete. Com isto, os histogramas das regiões centrais da imagem apresentam uma média de intensidade maior que os histogramas em regiões mais próximas dos limites da imagem e a adoção de um único limiar não produz bons resultados.

Uma melhor alternativa, portanto, são as técnicas locais que consistem em determinar um limiar para cada região estabelecida por uma janela. O tamanho da janela depende das características da imagem original. As técnicas locais são bastante conhecidas e existem vários algoritmos com diferentes respostas dependendo do tipo de imagem (BALLARD e BROWN, 1982; SEZGIN e SANKUR, 2004).

A definição de um limiar global pode ser feita interativamente a partir da análise do histograma da imagem ou usando métodos automáticos. Uma análise detalhada das técnicas de limiarização e sua eficiência pode ser encontrada em Sahoo et al (1988) e em Sezgin e Sankur (2004).

Dentre as técnicas existentes, uma das mais conhecidas é a de Otsu (1979), mas várias outras técnicas foram desenvolvidos recentemente. Para o contexto do presente trabalho, foram utilizadas as técnicas de Otsu (1979) e de Pun (1980).

A técnica de Otsu (1979) é baseada na análise discriminante e o valor do limiar é obtido supondo-se que os pixels da imagem podem ser classificados em duas classes (C_0 e C_1) que são o objeto e o fundo. Outro método global é o desenvolvido por Pun (1980) que é baseado na teoria da informação e, portanto, parte da premissa de que a geração de informação pode ser modelada como um processo probabilístico.

Uma das técnicas locais, considerada eficiente na avaliação realizada por Sezgin e Sankur (2004), é a de Niblack (1986). Entretanto, dependendo da aplicação analisada o método de Pun foi superior. Uma possibilidade similar é a de usar métodos globais, como os de Otsu e Pun, mas aplicá-los localmente, em pequenas janelas da imagem. Artero et al (2000) utilizaram uma técnica de limiarização local da imagem que consiste em determinar um limiar usando os *pixels* pertencentes a uma janela de 60 x 60 *pixels*. O valor de limiar assim obtido é utilizado para todos os *pixels* localizados na região central desta janela (uma região de tamanho 20 x 20 *pixels*). As dimensões destas janelas foram definidas empiricamente e podem ser alteradas dependendo das características da imagem. O limiar local pode ser definido por critérios como os de Otsu ou Pun, ou qualquer outro.

É importante ressaltar que muitos dos testes realizados por diferentes autores consideram um tipo específico de aplicação, por exemplo, documentos digitalizados, e a eficiência da técnica varia dependendo da aplicação e das condições nas quais as imagens foram tomadas. Como o foco do presente estudo está nas técnicas de correspondência e identificação de alvos, foram usadas as técnicas de Pun e Otsu, adaptadas para aplicação local (ver seção 3.3)

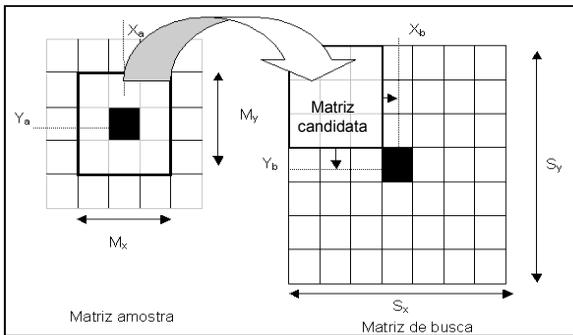
2.5 Classificação de alvos

A classificação ou identificação de alvos é um processo que consiste em rotular os objetos anteriormente segmentados em diferentes classes definidas *a priori*. Esta classificação pode ocorrer a partir da similaridade entre cores, níveis de cinza, textura etc. Segundo Jain et al (2000), as melhores abordagens para a classificação de padrões são: correspondência por padrão, classificação estatística, correlação estrutural ou sintática e redes neurais. Neste trabalho foi utilizado o método de correspondência por padrão, com algumas variantes, como a estratégia de reamostragem adaptativa dos padrões, que será vista na seção 3.

2.5.1 Correspondência por padrão (*Template matching*)

Segundo Schenk (1999), o método de correspondência por padrão consiste em comparar a distribuição de níveis de cinza de uma matriz amostra (padrão) com matrizes candidatas pertencentes a uma matriz de busca, usando uma função de correlação adequada (Figura 1). Várias são as funções para este tipo de correspondência: erro, erro quadrático, correlação, covariância cruzada e covariância cruzada modificada. A função mais usada na literatura e que proporciona melhores resultados é a covariância cruzada modificada ou coeficiente de correlação.

FIGURA 1 – Janelas de busca e pesquisa no processo de correspondência.
(Adaptado de ANDRADE, 1998).



Para a determinação da correspondência a matriz amostra (ou janela de referência) é comparada com cada matriz candidata na janela de referência, comparação esta feita com o cálculo de uma das funções especificadas anteriormente. O pixel que apresentar a melhor resposta é selecionado como ponto de correspondência.

A função covariância cruzada modificada, também chamada de coeficiente de correlação resulta da normalização da função covariância cruzada (Equação 3). O coeficiente de correlação, neste caso, varia de -1 a 1. O valor 1 para o coeficiente de correlação corresponde à medida de similaridade máxima, o valor 0 indica que não há correlação e o valor -1 indica correlação inversa.

$$\rho = \frac{\sigma_{rb}}{\sigma_r \sigma_b} = \frac{\sum (g_r(i, j) - \bar{g}_r)(g_b(i + a, j + b) - \bar{g}_b)}{\sqrt{\sum (g_r(i, j) - \bar{g}_r)^2 \sum (g_b(i + a, j + b) - \bar{g}_b)^2}} \quad (3)$$

Onde:

σ_{rb} - covariância entre a janela de referência (matriz amostra) e a matriz candidata;

σ_r - desvio-padrão da janela de referência;

σ_b - desvio-padrão da matriz candidata na janela de busca;

ρ - coeficiente de correlação;

g_r e g_b - níveis de cinza da janela de referência (matriz amostra) e da matriz candidata, respectivamente;

i, j - índices dos *pixels* na janela de referência e na matriz candidata;

a, b - valores para mudanças de posição em linha e coluna na janela de busca (matriz de busca);

\bar{g}_r, \bar{g}_b - médias dos níveis de cinza da janela de referência e da matriz candidata.

3. METODOLOGIA

Este trabalho tem como objetivo principal apresentar os resultados com as técnicas estudadas para a classificação ou identificação de padrões projetados com um sistema de luz estruturada. A classificação é um passo fundamental que permite a rotulação de cada alvo projetado. As etapas posteriores são o reconhecimento dos grupos de padrões, o cálculo das coordenadas dos centros de massa dos alvos ou dos cantos e, finalmente o cálculo das coordenadas 3D por intersecção. Apenas a etapa de classificação será discutida neste trabalho. É importante ressaltar que todos as técnicas descritas neste trabalho foram implementas em linguagem C++.

3.1 Geração dos Padrões

Para a realização da reconstrução 3D, o método de reconhecimento de padrões a ser usado necessita classificar cada padrão unicamente, pois cada ponto tem diferentes valores para as componentes do vetor diretor, usado para a determinação das coordenadas 3D.

Com base em estudos anteriores (SCALCO et al, 2000) e em outros trabalhos da literatura, chegou-se a várias possibilidades de abordagens, sendo a mais adequada, a utilização de padrões de luz branca com formas diferentes. Além disto, procurou-se gerar padrões de forma a privilegiar a existência de quinas, que poderiam ser usadas como pontos com localização precisa e que, conseqüentemente, garantiriam uma melhor densidade de reconstrução se comparados apenas com o centro de massa de cada padrão. Foram definidos cinco padrões elementares (Figura 2).

FIGURA 2 - Padrões primitivos.

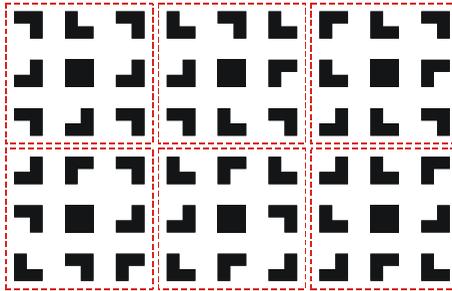


O padrão A (quadrado) foi definido como o pivô de uma estrutura bidimensional de vizinhança oito, pois foi o mais facilmente detectado pelo método de correspondência de imagens, nos testes realizados. Além disto, a geração dos

grupos de padrões seguiu algumas restrições, que eliminam a possibilidade de grupos de padrões ocorrerem a uma distância menor que 4 grupos.

O programa de geração automática de padrões tem como resultado a criação de um padrão composto pela combinação de cinco padrões primitivos. Estes padrões primitivos foram combinados em matrizes, definindo uma malha (Figura 3), que foi reproduzida por processo fotográfico, gerando-se um fotolito. O fotolito usado neste trabalho possui padrões com tamanho 9×9 *pixels* e com espaçamento de 7 *pixels* entre os padrões.

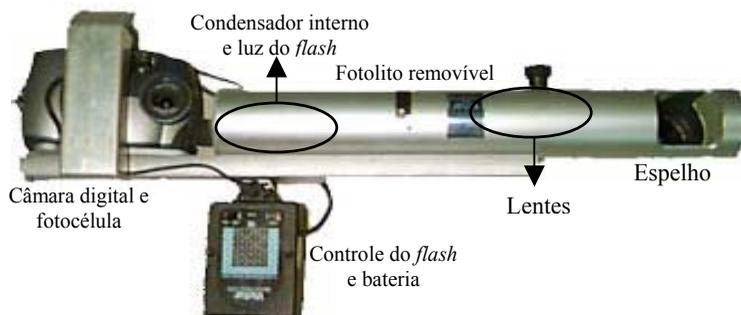
FIGURA 3 – Padrões primitivos combinados gerando estruturas.



3.2 Projeção e Coleta de Dados

O projetor de padrões é composto de uma câmara digital, fotocélula, condensador interno, fotolito, bateria, lentes e espelho (Figura 4). O fotolito foi colocado entre duas placas de vidro planas, usadas para absorver parte do calor gerado pela lâmpada.

FIGURA 4 - Sistema sensor e projetor.



O controlador do *flash* assegura a sincronização entre o instante de aquisição da imagem pela câmara e a iluminação da cena pelo projetor. Este tipo de iluminação instantânea evita o aquecimento do projetor e a deformação do padrão reproduzido no material fotográfico.

A câmara usada para a construção deste protótipo de sistema de luz estruturada foi a KODAK DX 3500, que possui uma distância focal nominal de 38 mm, resolução máxima de 1200 x 1800 pixels e tamanho do pixel de 19,44 μm (para um quadro focal de 35 mm).

No momento do disparo da câmara, a fotocélula capta a luz do *flash* da câmara e dispara o *flash* interno do projetor. A luz é direcionada pelo condensador para a objetiva e redirecionada pelo espelho, iluminando o objeto com os padrões. A imagem capturada pela câmara apresenta os padrões claros e o fundo escurecido. Com a identificação de cada padrão e a medição das coordenadas do centro de massa ou de seus cantos, pode-se reconstruir o objeto.

A seguir, são descritas as etapas de processamento de imagens aplicadas nas imagens coletadas pela câmara digital.

3.3 Aplicação do Método de Limiarização Local

Devido à distribuição irregular de luz pela imagem, realizou-se uma etapa prévia de limiarização local, para separar o fundo do sinal gerado pelos padrões projetados. Foi usado o método de limiarização local porque se verificou experimentalmente a necessidade de definir diferentes limiares para cada região da imagem.

Inicialmente, foi considerada a hipótese de definir um valor de limiar para cada *pixel* da imagem (calculado a partir de uma pequena região em torno do *pixel*).

Porém, devido ao grande volume de processamento que isto implicaria, uma nova estratégia foi desenvolvida e, resolveu-se implementar uma variante da limiarização pelos métodos de Otsu e de Pun, como definido por Artero et al (2000). Neste caso, o limiar é calculado a partir de uma subimagem $3w \times 3w$ e aplicado em uma janela $w_x \times w_y$. Isto permite uma transição suave de uma janela para outra, evitando o efeito de “ladrilho” na imagem limiarizada. *Pixels* com níveis de cinza inferiores ao limiar recebem zero enquanto os outros *pixels* permanecem com o nível de cinza original. Esta variante do método de limiarização não binariza a imagem, pois os tons de cinza do alvo serão necessários no processo de determinação precisa das coordenadas dos alvos.

3.4 Segmentação por Crescimento de Regiões

Para a classificação de padrões, propõe-se a etapa prévia de segmentação, pois isto evita que cada padrão seja comparado com toda a imagem, otimizando o processo, evitando diferentes respostas para o mesmo tipo de alvo, proporcionando resultados mais confiáveis e reduzindo o custo computacional.

A segmentação dos alvos na imagem é feita usando o método de crescimento de regiões. Esta técnica consiste no agrupamento e na rotulação de todos os *pixels* de cada padrão. Isto é feito percorrendo toda a imagem até encontrar um *pixel* que possua valor diferente de zero. Este critério pode ser usado, pois no passo de limiarização anterior, todos os *pixels* com tons de cinza abaixo do limiar local foram zerados. Este *pixel* é denominado de *pixel* semente e é rotulado. A partir deste *pixel* semente com coordenadas (i, j) , observam-se os *pixels* vizinhos: frente $(i, j+1)$, abaixo $(i+1, j)$ e trás $(i, j-1)$, que recebem o mesmo rótulo se atenderem à condição de agregação. Este processo é recursivo e termina quando os *pixels* vizinhos tiverem sido rotulados ou tiverem valor zero. Finalizada a segmentação do alvo, determinam-se as coordenadas do seu centro de massa e reinicia-se a varredura a partir do *pixel* semente do padrão segmentado anteriormente. O centro de massa é determinado a partir de uma média ponderada das discrepâncias entre os níveis de cinza dos *pixels* segmentados e o limiar, determinado na imagem limiarizada. Alvos pertencentes ao fundo recebem rótulo 1; zero indica que o *pixel* não foi rotulado e os padrões segmentados recebem rótulos maiores que 1. Este procedimento é realizado até que todos os *pixels* da imagem tenham sido rotulados. As informações geradas neste procedimento são: coordenadas do centro de massa de cada alvo segmentado; número do rótulo e coordenadas de todos os *pixels* pertencentes ao alvo; e, as coordenadas máxima e mínima de cada padrão, para a determinação da altura e da largura do alvo, visando uma posterior reamostragem.

3.5 Estratégias de classificação Baseadas na Correspondência por Padrão

Para verificar a eficiência do processo de correspondência por padrão, cinco funções de correlação foram implementadas em linguagem C++ e comparadas: erro, erro quadrático, covariância, correlação, covariância cruzada modificada. Os testes realizados mostraram que a função covariância cruzada modificada é a que apresenta melhores resultados, embora tenha um custo computacional maior, mas que não interfere no fluxo da aplicação. Na função de covariância cruzada modificada, a melhor correspondência será aquela em que o coeficiente de correlação for o mais próximo da unidade. Porém, se a imagem sofre rotações, variações na escala ou perda de parte do alvo, uma falsa correspondência pode ocorrer ou o coeficiente de correlação ser tão pequeno que não possa ser considerado. A partir de testes realizados, observou-se que um limiar para este coeficiente de correlação poderia ser 0,7, porque garantiria uma classificação satisfatória para as imagens usadas e não provocaria uma perda significativa dos alvos.

É importante ressaltar que a metodologia proposta neste trabalho visa comparar cada alvo previamente segmentado com o um conjunto de padrões pré-armazenados. Para isto, foram desenvolvidas três variantes para o processo de classificação dos alvos segmentados, que serão descritas nas seções seguintes.

3.5.1 Classificação por comparação dos alvos segmentados com padrões binários.

Esta técnica consiste na comparação dos alvos segmentados com os padrões pré-definidos binários e as suas instâncias armazenadas. Nesta abordagem os cinco padrões primitivos são instanciados e geram 45 padrões, que são armazenados para posterior comparação (Figura 5). Estes padrões foram gerados modificando-se empiricamente as escalas em x e y, a partir de observações realizadas na imagem original. O problema desta abordagem está no fato dos padrões serem binários e as bordas abruptas, o que não acontece nas imagens reais. Para tentar minimizar este problema foi definida uma segunda estratégia, mostrada na próxima seção.

Figura 5 - Padrões pré-definidos e as suas instâncias armazenadas.

Padrão \ Tamanho (pixels)	A	B	C	D	E
11 x 11					
13 x 13					
15 x 15					
17 x 17					
19 x 19					
13 x 15					
15 x 13					
15 x 17					
17 x 15					

3.5.2 Classificação por comparação dos alvos segmentados com padrões suavizados.

Neste caso, os mesmos 45 padrões (figura 5) são armazenados para posterior comparação. Porém estes padrões sofrem uma suavização pela média antes da comparação com os alvos segmentados, para simular a função de espalhamento que o dispositivo de aquisição causa na imagem. Isto faz com que as bordas dos padrões sejam ligeiramente borradas, o que as aproxima das imagens reais.

Estas duas metodologias (seções 3.5.1. e 3.5.2) têm alguns inconvenientes. Em primeiro lugar, é necessário armazenar um grande número de padrões e instâncias para obter correspondências aceitáveis. Mesmo armazenando muitos padrões, não é possível prever quais as variações de escala que ocorrerão na prática, devido à inclinação da superfície, e mesmo às variações de rotação. Quando a variação de escala é muito grande, a comparação pode falhar. Outro problema refere-se à posterior determinação das coordenadas dos cantos dos alvos com precisão subpixel, que é afetada pelas coordenadas aproximadas definidas na etapa de correspondência. Mesmo que a correspondência e identificação sejam corretas, se houver discrepâncias significativas entre o padrão e o alvo segmentado, o processo de determinação de coordenadas com precisão subpixel não será confiável. Este processo de determinação precisa de coordenadas dos alvos não será discutido neste trabalho. Por estas razões, definiu-se uma técnica que permite a geração adaptativa dos padrões, por reamostragem dos cinco padrões primitivos, que será descrita na seção seguinte (3.5.3).

3.5.3 Classificação por comparação dos alvos segmentados com padrões reamostrados adaptativamente.

Nesta técnica são armazenados apenas os 5 padrões primitivos, que são reamostrados adaptativamente, de acordo com as dimensões do alvo segmentado. Estas dimensões são obtidas a partir das discrepâncias entre as coordenadas máximas e mínimas em x e y, determinadas no processo de segmentação por crescimento de regiões. Estas coordenadas permitem determinar a largura e altura do alvo segmentado. Com as dimensões do alvo conhecidas, os 5 padrões primitivos são reamostrados, utilizando a interpolação bilinear. Este procedimento pode ser usado porque as imagens sofrem uma pequena rotação que pode ser desconsiderada. As vantagens deste método em relação aos dois primeiros estão relacionadas ao número de padrões armazenados e a melhor determinação das coordenadas aproximadas para a localização precisa posterior.

Um problema está na definição das dimensões da matriz de amostra a ser armazenada. Para definir melhor estas dimensões foram realizados alguns experimentos que serão apresentados no capítulo 4.

4 EXPERIMENTOS

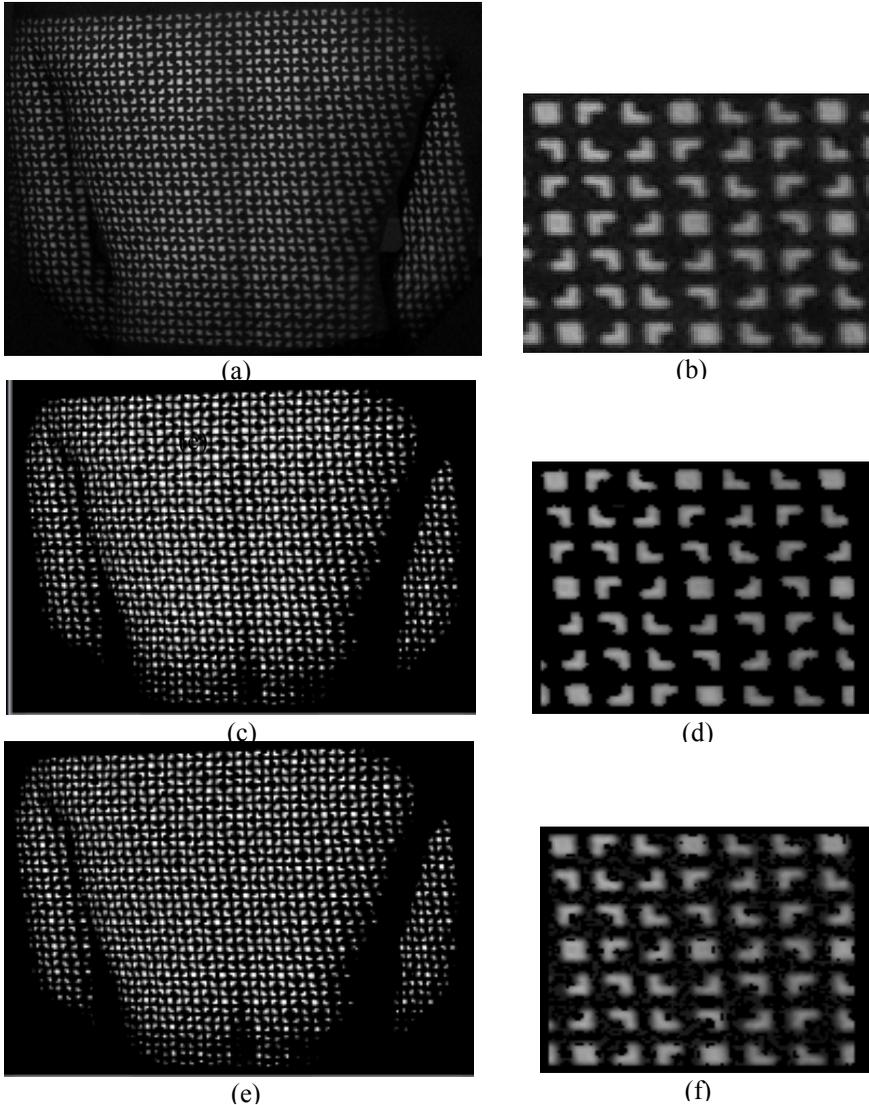
Os experimentos foram realizados com imagens coletadas com o sistema em condições regulares de operação. O objetivo destes experimentos é avaliar o comportamento das diferentes estratégias de classificação, bem como de processos prévios, como a limiarização. Foram usadas porções das imagens originais, para permitir uma análise mais detalhada dos resultados. Foram feitos experimentos com diferentes técnicas de realce e limiarização, com a segmentação por crescimento de regiões e com diferentes estratégias de classificação. O objetivo dos experimentos não foi o de explorar exaustivamente todas as possibilidades das diferentes técnicas nas várias fases de processamento, mas o de encontrar um caminho que proporcionasse resultados aceitáveis. Um critério empírico foi considerar aceitável uma técnica que proporcionasse uma taxa de classificação correta acima de 90%. Apenas como comparação, sabe-se que os algoritmos de correlação usados em sistemas fotogramétricos têm uma taxa de acerto ao redor de 85%.

4.1 Experimentos com as Técnicas de Limiarização Local

Como foi discutido anteriormente, a etapa de limiarização permite separar os alvos projetados do fundo. Devido à variação na iluminação, os métodos globais não funcionam adequadamente, eliminando alvos na periferia da imagem. Foram então feitos testes com duas possibilidades de processamento local, que serão descritos a seguir. O objetivo dos testes foi o de avaliar as técnicas de limiarização

local usando os operadores de limiarização automática de Otsu e Pun, como descrito nas seções 2.4 e 3.3.

FIGURA 5 – Limiarização local (a) Imagem original (tomada de um dorso humano) com o sistema de luz estruturada; (b) Parte da imagem original; (c) (d) Limiarização com o método de Otsu local; (e) (f) Limiarização com Pun local.



No método de limiarização local empregado, define-se o tamanho da janela a ser percorrida na imagem e o método de cálculo do limiar. Para cada percurso da janela na imagem, um limiar é determinado a partir dos métodos de Otsu ou Pun e aplicado na região central da janela. Os *pixels* com valores abaixo do limiar determinado pelo método de Otsu ou de Pun recebem nível de cinza zero e os *pixels* com níveis de cinza superiores a este valor, permanecem com o mesmo valor de tom de cinza.

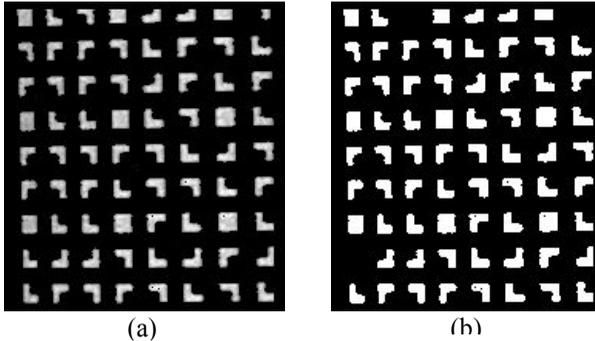
Os resultados deste processo podem ser observados na Figura 5. Como se podem perceber visualmente, ambos os métodos de limiarização conseguiram eliminar completamente os ruídos do fundo, sem prejudicar o processo de reconhecimento. Além disto, foram comparados três diferentes tamanhos de janelas (3x3, 5x5 e 7x7). A partir de testes, observou-se que a janela 3x3, usando o método de Otsu, proporcionou melhores resultados aliados a um menor custo computacional. Como exemplo, para uma janela 3x3, significa que o limiar foi calculado em uma janela 9x9, mas aplicado apenas à porção central (3x3).

Verifica-se que esta estratégia local é mais adequada, pois a aplicação de um único valor de limiar para toda a imagem, eliminaria os alvos na periferia da imagem. Outros critérios definidos na literatura para definição do valor local de limiar poderiam ser empregados, mas pode-se afirmar que o método de Otsu, neste caso foi suficiente.

4.2 Experimento com o Processo de Segmentação por Crescimento de Regiões

O objetivo deste experimento foi verificar a eficiência do processo de segmentação por crescimento de regiões nas imagens do sistema de luz estruturada. Após o processo de limiarização, todos os pixels do fundo foram marcados com o valor zero e os demais mantiveram seus tons de cinza originais. O processo de segmentação agrega os pixels similares e atribui um rótulo a cada grupo. Grupos com um número de pixels pequeno, são eliminados, entendendo-se que se constituem em alvos incompletos ou em ruído. O resultado deste processo pode ser observado na Figura 6.

FIGURA 6 – Segmentação de imagens. (a) Imagem limiarizada;
(b) Imagem segmentada.



Como se pode observar na Figura 6b, alguns alvos não foram segmentados. Isto ocorre devido às restrições quanto ao tamanho do alvo. Esta predição é possível com base na relação da distância focal da câmara com as distâncias máxima e mínima do centro perspectivo da câmara à superfície do objeto. Assim, alvos com quantidade de *pixels* não pertencentes ao intervalo estabelecido são eliminados após a segmentação. Esta restrição ocorre para não segmentar grupos de *pixels* que sejam ruídos na imagem ou alvos muito grandes, que são consequência de alvos pertencentes à borda do objeto na cena ou junção de alvos, proporcionando uma classificação equivocada.

4.3 Experimentos com o Processo de Classificação

A partir da imagem segmentada, inicia-se uma das etapas fundamentais para a realização da reconstrução: a classificação dos alvos. Esta etapa é importante porque os seus resultados interferem na localização precisa e posterior reconstrução pois as coordenadas aproximadas dos cantos dos alvos são determinadas nesta etapa. A classificação de alvos objetiva informar a melhor correspondência do alvo segmentado quando comparado com os padrões armazenados. Serão apresentados experimentos com as três estratégias definidas na seção 3.5.

4.3.1 Experimentos com a técnica de classificação por comparação dos alvos segmentados com padrões binários.

Neste método são armazenadas as imagens dos cinco padrões primitivos e suas respectivas instâncias. Também são armazenadas as coordenadas em *pixels* dos cantos, que podem ser usadas como valores aproximados no processo posterior de localização precisa. O resultado deste processo de classificação pode ser visto na

Figura 7. Na Figura 7b, os alvos encontram-se sem rotação porque se concentram na parte central do cilindro, enquanto que na Figura 7d, os alvos estão distorcidos pela inclinação da superfície.

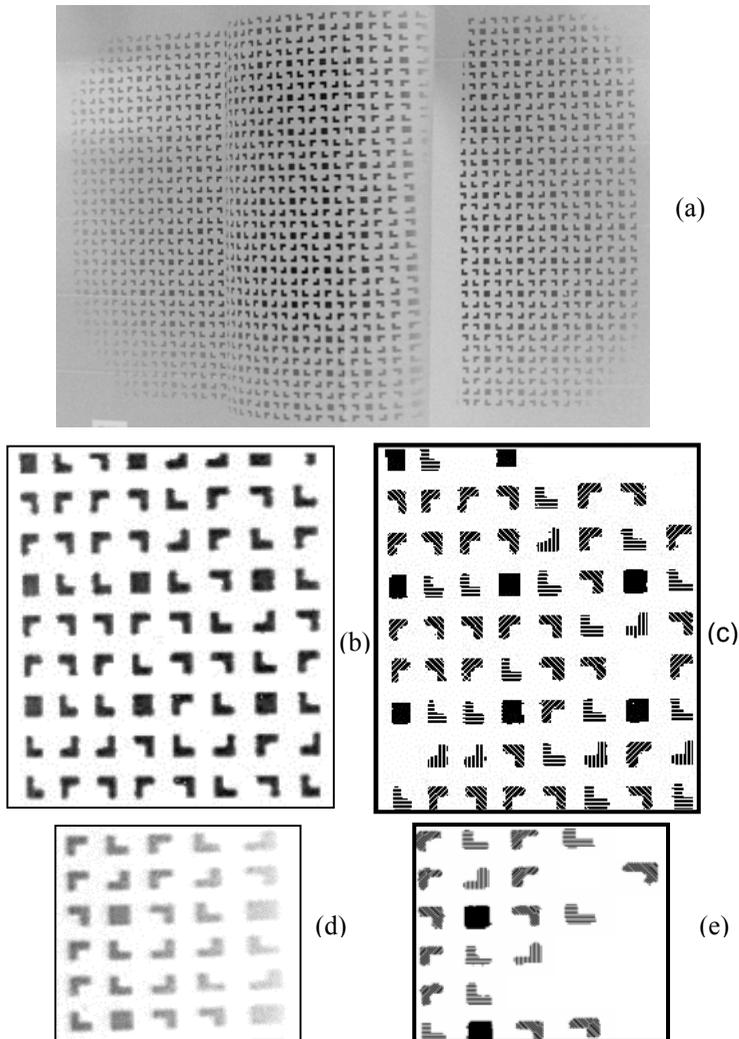
Por uma questão de visualização, foram inseridas texturas nos alvos classificados: padrão A (preto); padrão B (retas horizontais); padrão C (retas verticais); padrão D (retas oblíquas-45°) e padrão E (retas oblíquas-135°).

No experimento mostrado nas figuras 7b e 7c todos os alvos foram classificados corretamente. Os alvos que não foram identificados apresentavam-se parcialmente incompletos ou deformados pela superfície e não geraram um coeficiente de correlação maior que o pré-definido como limiar (0,7) com nenhum dos padrões ou suas instâncias.

Uma dificuldade observada foi a identificação de alvos em escalas diferentes em x e y e distorcidos (Figura 7d), já que não é recomendável gerar muitos padrões instanciados, porque isto aumentaria o custo computacional. Nestes casos, como os alvos sofrem modificações diferentes em escalas x e y, mesmo que ocorra uma identificação correta, as coordenadas aproximadas não são confiáveis.

FIGURA 7 – Resultados da correspondência por correlação usando padrões binários (os tons de cinza das imagens estão invertidos):

- (a) Imagem dos alvos projetados em um cilindro;
- (b) Imagem realçada de uma parte da imagem sem rotação;
- (c) Classificação dos alvos para a imagem com padrões projetados distorcidos;
- (d) Imagem realçada de uma imagem com variação na rotação;
- (e) Classificação dos alvos para a imagem com padrões projetados distorcidos.



Vale ressaltar que este procedimento proporciona bons resultados porque é pequena a percentagem de alvos não reconhecidos (em torno de 15%). Isto garante um grande número de vértices para a localização precisa o que permite uma melhor

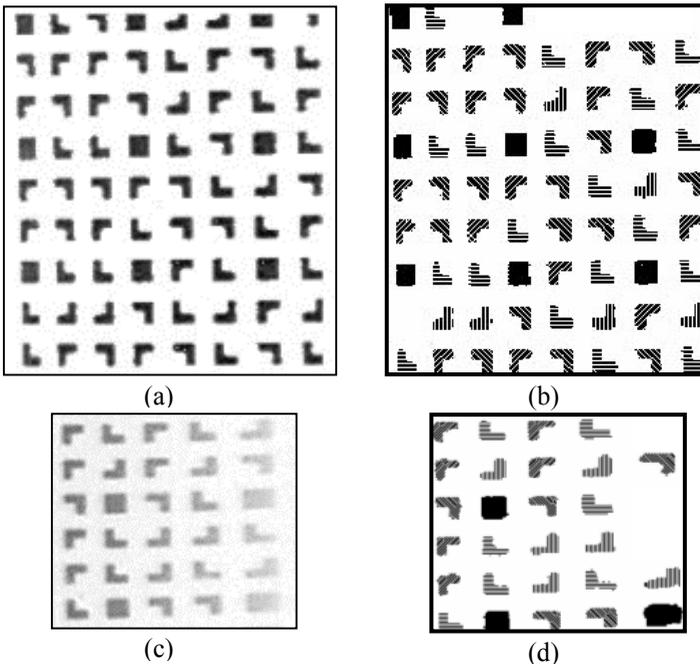
reconstrução do objeto se comparada com a reconstrução obtida apenas pelo centro de massa de cada alvo.

4.3.2 Experimentos com a técnica de classificação por comparação dos alvos segmentados com padrões suavizados.

Como pode ser observado na Figura 8, a correspondência por correlação usando padrões armazenados suavizados garantiu melhores resultados que o método usando imagens binárias, mostrado na seção anterior. Em imagens com pequena rotação, este método conseguiu classificar aproximadamente 94% dos alvos segmentados enquanto em imagens com variação na rotação, 90% dos alvos foram classificados corretamente.

FIGURA 8 – Resultados da correspondência por correlação usando padrões suavizados pela média (os tons de cinza estão invertidos)

- (a) Imagem realçada de uma imagem sem variação de rotação;
- (b) Classificação dos alvos para a imagem sem variação da rotação;
- (c) Imagem realçada de uma imagem com padrões projetados distorcidos;
- (d) Classificação dos alvos para a imagem com padrões projetados distorcidos.



Este melhor resultado se deve à melhor correspondência entre as imagens reais dos alvos com os padrões suavizados.

Mesmo considerando que os resultados anteriores eram satisfatórios, verificou-se a possibilidade de refinar ainda mais os resultados, usando a técnica de reamostragem adaptativa, descrita na seção 3.5.3..

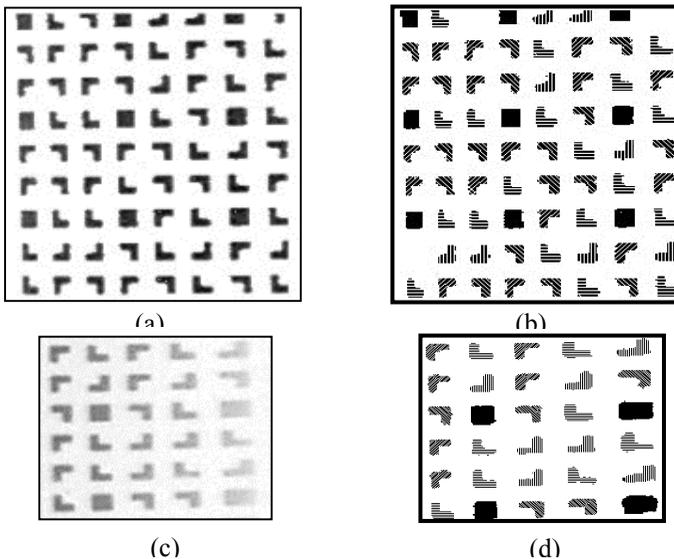
4.3.3 Experimentos com a técnica de classificação por comparação dos alvos segmentados com padrões reamostrados adaptativamente.

Esta técnica consiste em armazenar apenas os cinco padrões primitivos e a partir do conhecimento da dimensão dos alvos segmentados, reamostrá-los em tempo real usando o método de interpolação bilinear. Neste caso, três experimentos foram realizados: reamostrando-se padrões primitivos com dimensão 23×23 pixels, 90×90 pixels (sem fundo) e 90×90 pixels (com fundo). Foram testados diferentes tamanhos de padrões para verificar o efeito da reamostragem no processo de correspondência.

1º Experimento: Padrões primitivos com dimensão 23×23 pixels

FIGURA 9 – Resultados da correspondência por correlação usando padrões reamostrados (Padrões 23×23 pixels)

- (a) Imagem realçada sem variação de rotação;
- (b) Classificação dos alvos para a imagem sem variação da rotação;
- (c) Imagem realçada com padrões projetados distorcidos;
- (d) Classificação dos alvos para a imagem com padrões projetados distorcidos.



Esta dimensão dos padrões armazenados de 23x23 foi escolhida empiricamente, a partir de observações realizadas na imagem. A dimensão dos alvos segmentados foi obtida a partir da diferença entre as coordenadas máximas e mínimas de cada alvo. Este procedimento pode ser usado, neste caso, porque as imagens usadas possuem uma pequena variação da rotação. Com esta técnica, aproximadamente 96% dos alvos foram classificados corretamente para imagens com pequena rotação e em mesmo em algumas regiões da imagem com padrões projetados distorcidos, todos os alvos foram classificados (Figura 9). É importante ressaltar que este procedimento conseguiu classificar corretamente mesmo alguns alvos incompletos.

Posteriormente, observou-se, que esta técnica de classificação poderia ser melhorada ainda mais a partir do uso de padrões com dimensão proporcional ao do alvo projetado ao invés da escolha empírica de sua dimensão, com base apenas na observação dos alvos na imagem. Neste caso, a dimensão escolhida para os padrões armazenados foi de 90 x 90 *pixels*.

2º Experimento: Padrões primitivos com dimensão 90 x 90 *pixels* (reamostragem sem o fundo)

Neste experimento, foi reamostrada apenas a parte branca do alvo e posteriormente, foi inserido o fundo preto com dimensão de 2 *pixels* (Figura 10). Neste experimento, 85% dos alvos segmentados foram classificados e, dentre estes, todos os alvos foram identificados corretamente. Os resultados obtidos com este método podem ser observados na Figura 11.

FIGURA 10 – Reamostragem dos padrões primitivos sem o fundp.
Padrão primitivo com 90x90 pixels;
Padrão primitivo sem o fundo;
Padrão reamostrado de acordo com o tamanho do alvo
Reinserção do fundo com 2 pixels de largura.

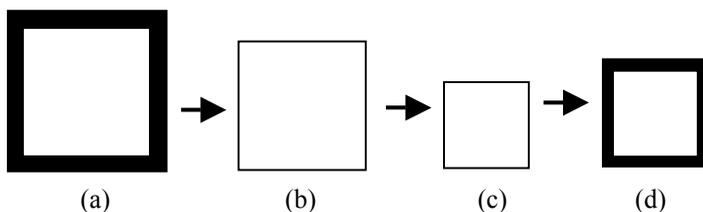
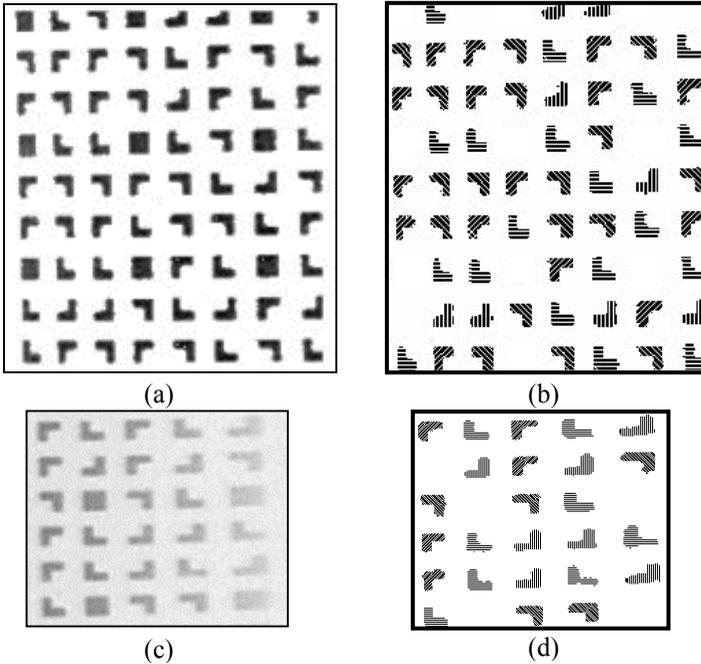


FIGURA 11 – Resultados da correspondência por correlação usando padrões reamostrados sem bordas (Padrões $90 \times 90 \text{ pixels}$)

(a) Imagem realçada de uma imagem sem variação de rotação;

(b) Classificação dos alvos para a imagem sem variação da rotação;



(c) Imagem realçada de uma imagem com padrões projetados distorcidos;

(d) Classificação dos alvos para a imagem com padrões projetados distorcidos.

3º Experimento: Padrões primitivos com dimensão $90 \times 90 \text{ pixels}$ (reamostragem com o fundo)

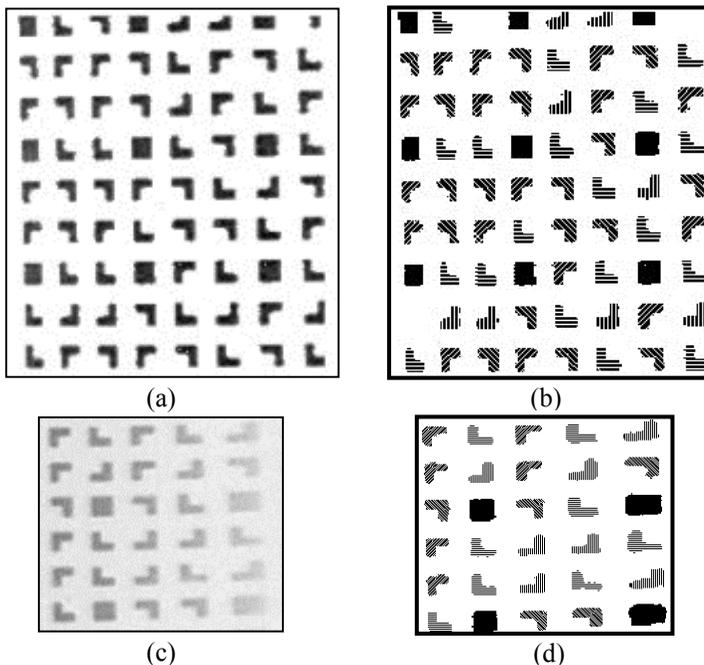
Como a técnica anterior (Figuras 10e 11) não proporcionou os resultados esperados, decidiu-se usar padrões com dimensão proporcional ao do alvo projetado ($90 \times 90 \text{ pixels}$), mas, ao invés de reamostrar o alvo sem o fundo, este foi reamostrado incluindo a moldura de dimensão 2 pixels . Os resultados desta técnica podem ser observados na Figura 12, mostrando uma taxa de acertos maior que no experimento 2. Pode-se observar, ainda, que na imagem com os alvos distorcidos

pela inclinação da superfície (figura 12c) todos os alvos foram segmentados e corretamente classificados.

A partir destes experimentos realizados pode-se observar que as técnicas usando a reamostragem por interpolação bilinear em tempo real proporcionaram melhores resultados que a comparação com padrões estáticos pré-armazenados. Além disso, usando o método de reamostragem por interpolação bilinear, pode-se diminuir o limiar para valores como 0,5, e mesmo assim o procedimento ainda consegue correlacionar perfeitamente sem muitas perdas de alvos.

FIGURA 12 – Resultados da correspondência por correlação usando padrões reamostrados incluindo o fundo (Padrões 90 x 90 *pixels*)

- (a) Imagem realçada de uma imagem sem variação de rotação;
- (b) Classificação dos alvos para a imagem sem variação da rotação;
- (c) Imagem realçada de uma imagem com padrões projetados distorcidos;
- (d) Classificação dos alvos para a imagem com padrões projetados distorcidos.



Verificou-se, ainda, que não é viável reamostrar apenas a imagem sem as molduras e depois inseri-las, como foi mostrado no experimento2, da seção 4.3.2. Este procedimento faz com que a imagem não sofra uma suavização e os valores dos coeficientes de correlação são inferiores aos valores estipulados anteriormente. Realizando este mesmo procedimento com os alvos pré-definidos, a perda de alvos é mais elevada e, ainda, alguns alvos já começam a ser correlacionados erroneamente. Por outro lado, o método de classificação usando alvos pré-definidos também proporcionou bons resultados e poderia ser melhorado ainda mais, se o número de padrões armazenados fosse maior. Porém, isso não é viável porque aumentaria muito o custo computacional

5 CONSIDERAÇÕES FINAIS

O processo de classificação de alvos por correspondência de padrões, apresentado neste trabalho, produziu bons resultados em imagens coletadas com um sistema de luz estruturada. Com relação ao tipo de padrões pré-definidos a serem usados (*template*), notou-se que o uso de imagens suavizadas proporcionou melhores resultados do que padrões binários. Além disso, a técnica usando padrões reamostrados em tempo real (comparação adaptativa) foi o que apresentou os melhores resultados.

Outras estratégias que podem ser usadas para melhorar os resultados são: Utilizar o conhecimento *a priori* a respeito das posições e das orientações da câmara e do projetor em relação ao espaço objeto para fazer a predição de localização e aspecto do alvo;

Realizar uma predição da rotação e da escala do alvo a partir do conhecimento das distâncias e dos ângulos entre os alvos segmentados e a partir disto reamostrar adaptativamente os alvos primitivos. Esta estratégia parece ser confiável, porém, possui problemas em áreas de descontinuidades;

Realizar uma operação de melhoramento local da imagem por meio da técnica de especificação do histograma, o que deve melhorar a localização precisa dos alvos;

Melhorar o sistema de projeção e captura com a aquisição de uma nova câmara de melhor resolução e construção mais rígida.

REFERÊNCIAS BIBLIOGRÁFICAS

- ANDRADE, J. B. Fotogrametria. Curitiba. SBEE, 1998.
- ARTERO, A. O.; TOMMASELLI, A. M. G. Limiarização Automática de Imagens Digitais. *Boletim de Ciências Geodésicas, Curitiba*, Pr., v. 6, n. 1, p. 38-48, 2000.

- BALLARD, D.H.; BROWN, C.M. *Computer Vision*, New Jersey: Prentice-Hall, 1982, 522p.
- BATTLE, J.; MOUADIBB, E.; SALVI, J. Recent progress in coded structured light as a technique to solve the correspondence problem: a survey. *Pattern Recognition*. v. 31, n. 7, p. 963-982, 1996.
- EL-HAKIM, S. F. Vision-based automated measurement techniques. In: ATKINSON, K.B. *Close Range Photogrammetry and Machine Vision. Department of Photogrammetry and Surveying*, University College London. Whittles Publishing, 1996.
- GONZALEZ, R. C.; WOODS, R. E. *Processamento de Imagens Digitais*, Tradução: Roberto Marcondes César Junior e Luciano da Fontoura Costa, Ed. Edgard Blücher, São Paulo, 509p., 2000.
- JAIN, A. K. *Fundamentals of Digital Image Processing*. Prentice Hall, 569p., 1989
- JAIN, A. K.; DUIN, R.P.W.; MAO, J. Statistical Pattern Recognition: A Review. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. v. 22, n. 1, p. 264-323, 2000.
- NEWTON, I.; MITCHELL, H.L. Medical Photogrammetry. In: ATKINSON, K. B. *Close Range Photogrammetry and Machine Vision. Department of Photogrammetry and Surveying*, University College London. Whittles Publishing, 1996.
- OTSU, N. A threshold selection method from gray level histograms. *IEEE Transactions on Systems, Man and Cybernetics*. v.9, p.62-66, 1979.
- PUN, T. A new method for gray-level picture threshold using the entropy of the histogram. *Signal Processing*. v. 2, n.3, p. 223-237, 1980.
- SAHOO, P. K.; SOLTANI, S.; WONG, A. K. C., An Survey of Thresholding Techniques, *Computer Vision, Graphics and Image Processing*, n. 41, p. 233 – 260, 1988.
- SCALCO, P. A. P.; TOMMASELLI, A. M. G. Determinação automática de correspondência em um sistema de reconstrução com luz estruturada.. *Boletim de Ciências Geodésicas*, Curitiba, Pr., v. 7, n. 2, p. 41-62, 2001.
- SCHENK, T. Digital Photogrammetry. *TerraScience*, v.1, 1999.
- SEZGIN, M.; SANKUR, B. Survey over image thresholding techniques and quantitative performance evaluation. *Journal of Electronic Imaging*, v. 13, n.1, p. 146-165, 2004.
- SONKA, M.; VAKLAV, H; BOYLE, R. *Image Processing, Analysis, and Machine Vision*. PWS PUBLISHING. 770p, 1998.

AGRADECIMENTOS

Os autores deste artigo agradecem à Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) pelo apoio ao projeto, através de bolsas de mestrado e doutorado. Agradecem, ainda, à empresa AVR Instrumental, de Curitiba, pela adaptação e montagem do projetor e aos revisores pelas valiosas sugestões.

(Recebido em setembro/04. Aceito em fevereiro/05.)