Scanner 3D

Projecto da Disciplina de Visão

12 de Março de 2002

1 Introdução

Uma das funções mais valiosas do nosso sistema visual consiste na obtenção da estrutura tridimensional do mundo envolvente, em particular da forma dos objectos com que lidamos no dia–a–dia. A disponibilidade de ferramentas de computação cada vez mais rápidas, o progresso da computação gráfica e a expansão global da *World Wide Web* têm gerado muito interesse em sistemas de percepção da geometria dos objectos. As aplicações são numerosas e vão desde a animação e entretenimento, *design* industrial, robótica e realidade virtual, etc...

Neste trabalho pretende-se desenvolver um sistema prático e simples de obter a forma de objectos, utilizando um conjunto mínimo de equipamento, para além de um computador e uma câmara: uma lâmpada de secretária, um lápis e outro objecto rectilíneo, um tabuleiro de xadrez – equipamento barato e disponível em qualquer lar. O sistema a desenvolver é baseado no princípio de *luz estruturada*, ou seja, na projecção de padrões de iluminação sobre o objecto que permitam obter a reconstrução da sua forma mesmo em zonas que não contenham textura natural.

2 Descrição da ideia

Observe a figura 1. Uma câmara observa uma cena iluminada por uma fonte de luz aproximadamente

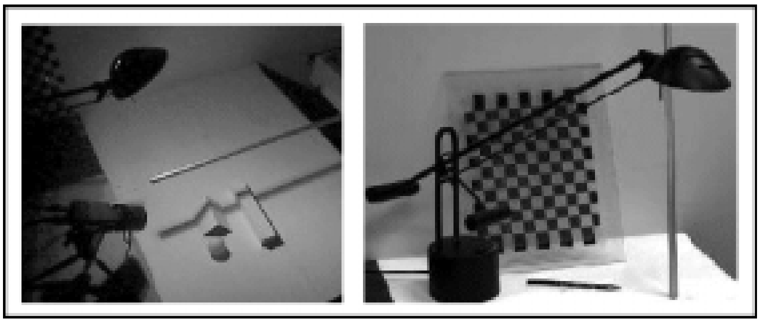

Figura 1: O *setup* experimental do sistema proposto

pontual (p.ex. uma lâmpada de secretária). A cena consiste num plano (a secretária) contendo objectos em cima. O utilizador projecta uma sombra na cena, interpondo entre os objectos e a fonte de luz um objecto rectilíneo (p.ex. um lápis), definindo assim um plano de sombra.

A câmara adquire uma sequência de imagens enquanto o plano de sombra "varre"a totalidade do objecto. A informação contida no contorno de sombra deformado pela superficie do objecto, vai nos permitir a recuperação da sua forma (ver figura 2). O objectivo final é obter a profundidade da cena em todos os *pixels* da imagem.

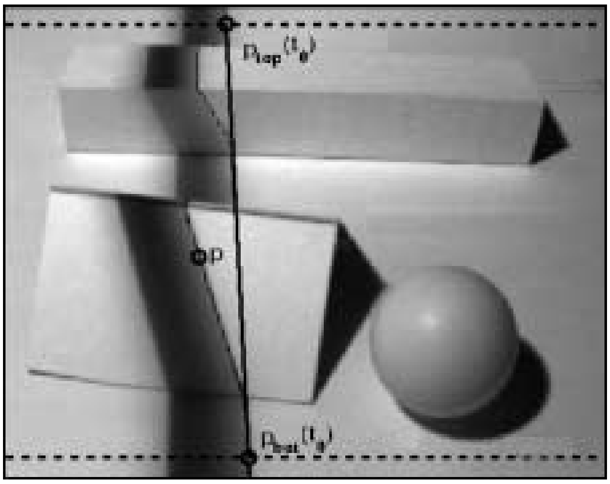

Figura 2: Imagem observada pela câmara

3 A Geometria do Problema

Observe atentamente a figura 3. Temos como objectivo estimar a posição 3D de cada ponto (P) observável sobre o objecto.

- Considera-se uma fonte de luz pontual, localizada na posição L.
- Denota-se por π_s o plano de sombra projectado na cena e por π_d o plano da secretária.
- Admite-se que duas porções da sombra projectada sobre a secretária são visíveis (linhas de referência superior e inferior – ver também figura 2).
- A intersecção do contorno de sombra com essas linhas define os pontos p_{top} e p_{bot}.
- Os pontos correspondentes sobre a secretária (P_{top} e P_{bot}) podem ser estimados pela intersecção dos raios ópticos (O, p_{top}) e (O, p_{bot}) com o plano da secretária.
- O plano de sombra π_s fica então definido pelos conjunto de pontos 3D $(P_{\text{top}}, P_{\text{bot}}, L)$.
- Finalmente, para todos os pontos p da imagem pertencendo ao contorno de sombra, pode-se calcular a profundidade do ponto correspondente sobre o objecto P, efectuando a intersecção do raio óptico (O, p) com o plano de sombra π_s.

Nesta descrição admitiu-se que:

- os parâmetros intrínsecos da câmara são conhecidos (obtidos numa fase prévia de calibração da câmara);
- a posição relativa entre a câmara e o plano da secretária é conhecida (obtida numa fase prévia de estimação da pose da câmara);
- a posição relativa da fonte de luz é conhecida (obtida numa fase prévia de estimação da posição da fonte de luz).

Com excepção da calibração da câmara, os outros pontos serão também abordados neste trabalho. Juntamente com uma descrição mais formal da fase de estimação de forma do objecto, as fases de estimação de pose da câmara e de estimação da posição da fonte de luz são descritas nas secções seguintes.

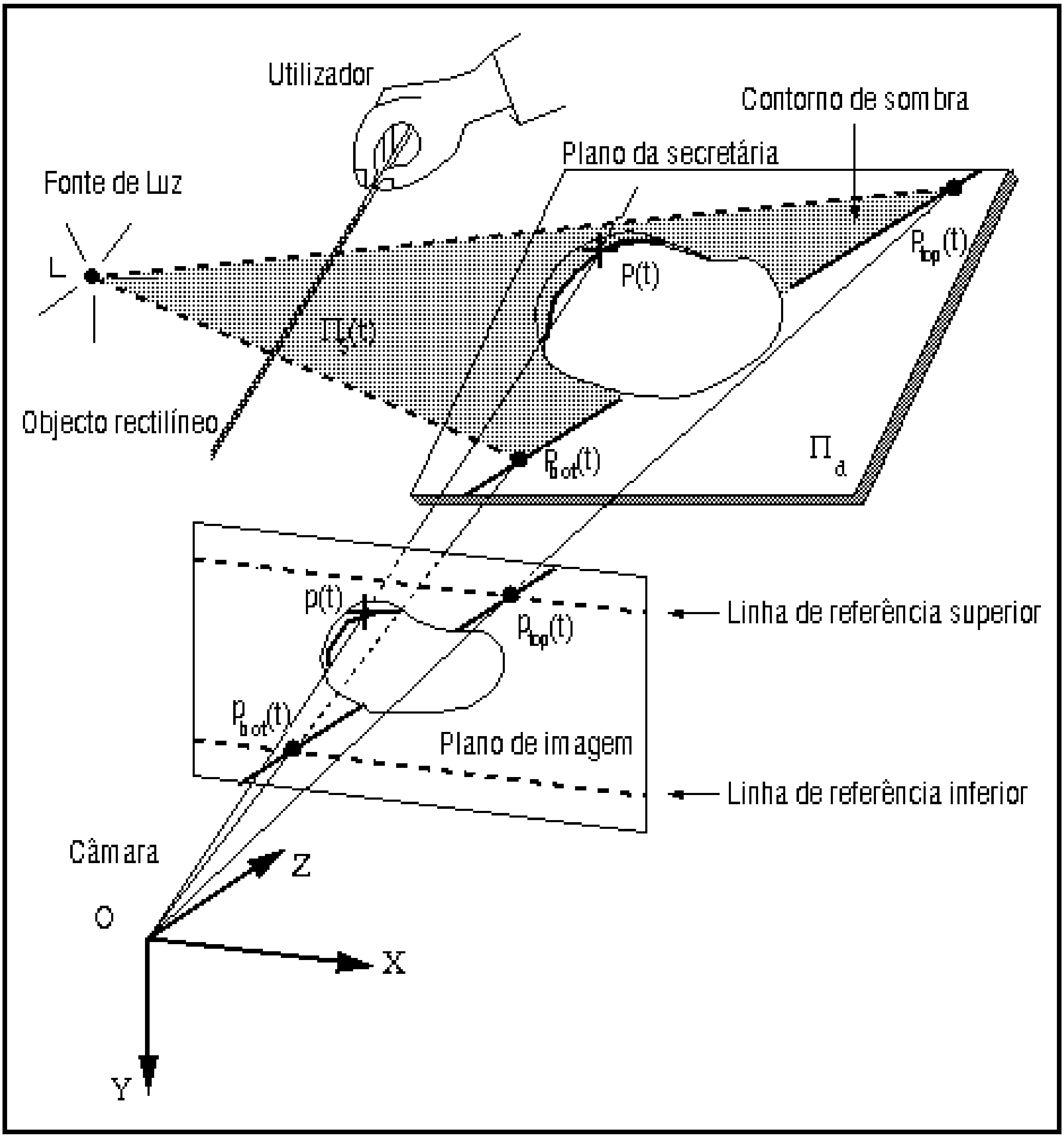

Figura 3: Representação geométrica do problema

4 Estimação da pose da câmara

Observe a figura 4. Pretende-se obter a transformação de coordenadas que relaciona a câmara com o

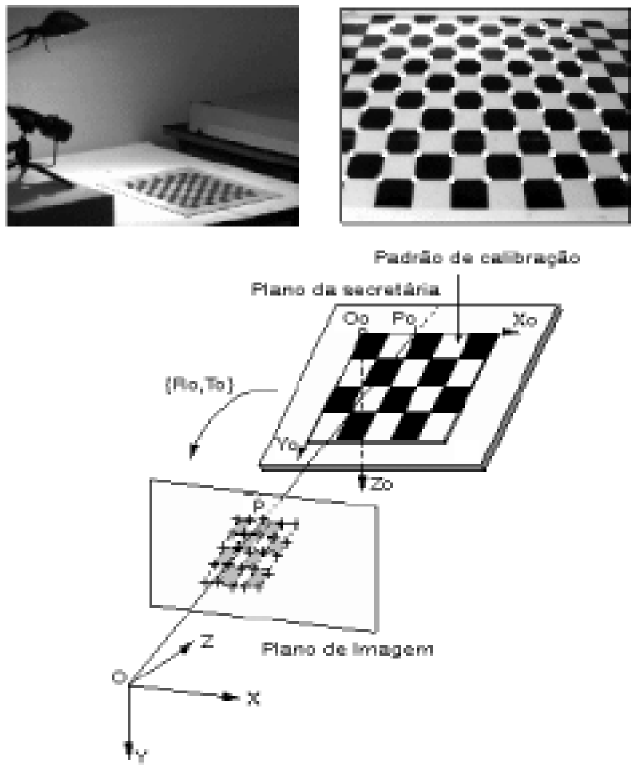

Figura 4: Estimação da pose relativa câmara–secretária

plano da secretária. Para isso vamos utilizar uma grelha de calibração conhecida em cima da secretária – um tabuleiro de xadrez.

Seja $\mathcal{D} = \{O_0, X_0, Y_0, Z_0\}$ o sistema de coordenadas associado à grelha de calibração e $\mathcal{C} = \{O, X, Y, Z\}$ o referencial da câmara. São conhecidos N pontos da grelha, P_0^n, $n \in [1, \ldots, N]$, que correspondem aos cantos das casas do tabuleiro.

A transformação de coordenadas que relaciona os dois referenciais pode ser expressa em coordenadas homogéneas por uma matriz T_0:

$$T_0 = \begin{bmatrix} R_0 & t_0 \\ 0 & 1 \end{bmatrix}, \; R_0 = \begin{bmatrix} r11 & r12 & r13 \\ r21 & r22 & r23 \\ r31 & r32 & r33 \end{bmatrix}, \; t_0 = \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix}$$

Pontos $P_0 = [X_0, Y_0, Z_0, 1]^T$ no referencial da grelha e pontos $P = [X, Y, Z, 1]^T$ no referencial da câmara, estão relacionados da seguinte forma:

$$P = T_0 \cdot P_0$$

Dada uma estimativa inicial para a transformação que relaciona os dois sistemas de coordenadas, T_0^i, pode-se tentar prever a posição da projecção dos pontos da grelha no plano de imagem:

$$\bar{p}^n = \Pi \left(T_0^i \cdot P_0^n \right)$$

em que Π corresponde à função de projecção associada ao modelo de formação de imagem adoptado. No trabalho presente vamos utilizar a projecção perspectiva [2].

Dada esta estimativa inicial para T_0, de um modo geral os pontos previstos $\bar{p}_n$ não correspondem aos observados p_n (a sua posição real pode ser medida nas imagens adquiridas, por um processo manual ou automático). O processo de estimação consiste em procurar o valor de T_0 que minimize a diferença quadrática média entre os pontos previstos e os pontos observados:

$$T_0 = \arg\min_{T_0} \sum_{n=1}^{N} \|p^n - \bar{p}^n\|^2 \quad (1)$$

O procedimento de minimização pode ser efectuado recorrendo a rotinas do MATLAB (ver secção 6). Obtida a matriz de transformação T_0, é possível obter as coordenadas 3D de qualquer ponto projectado na imagem que pertença ao plano da secretária.

5 Estimação da posição da fonte de luz

Observe a figura 5. Pretende-se localizar a posição da fonte de luz no espaço (L). Para isso utilizamos

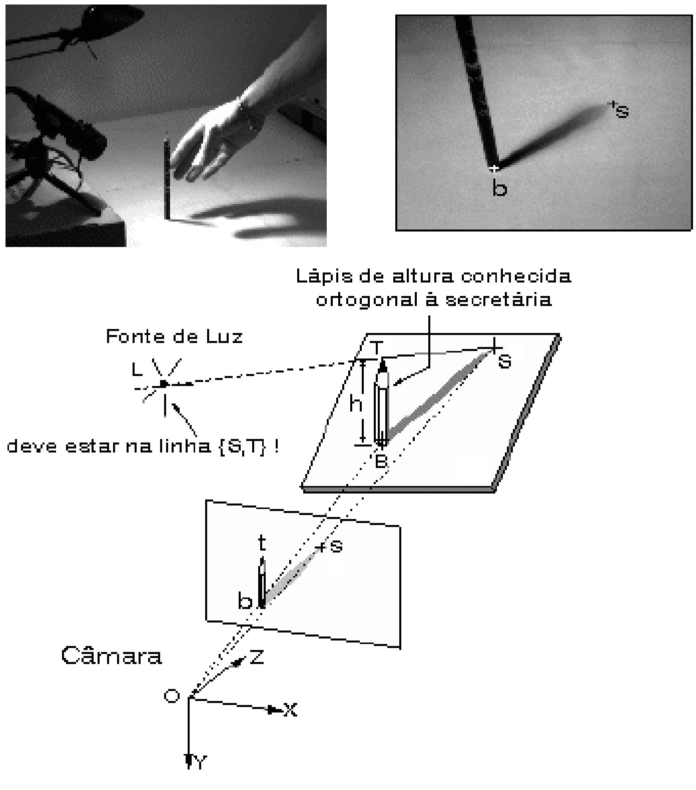

Figura 5: Estimação da posição da fonte de luz

um objecto com altura conhecida h e que se consiga colocar ortogonalmente ao plano da secretária (p.ex. um lápis). Vamos coleccionar M imagens da sombra formada pelo lápis na secretária. Da localização da sombra podemos obter, manual ou automaticamente:

- a projecção na imagem da base do lápis (b^m);
- a projecção na imagem da sombra da extremidade superior do lápis (s^m);

com $m \in [1, \ldots, M]$.

Como estes pontos se encontram sobre a secretária é fácil calcular as suas posições 3D (B^m e S^m) – basta intersectar os raios ópticos $\{O, b^m\}$ e $\{O, s^m\}$ com o plano da secretária. Embora a extremidade superior do lápis não se encontre sobre o plano da secretária, é ainda assim possível obter a sua posição 3D (T^m), uma vez que o lápis é ortogonal ao plano da secretária.

Para cada situação m, a fonte de luz deve-se encontrar algures sobre a linha que passa por S^m e T^m, que denominaremos de Δ^m. Numa situação ideal, seriam suficientes duas medições e a posição da fonte de luz seria determinada pela intersecção das rectas Δ^1 e Δ^2. No entanto, em condições normais, as rectas Δ^m não se intersectam devido a ruído e erros de medição, e procede-se à minimização de um funcional, como se descreve de seguida:

- Pontos $\bar{L}^m$ sobre a linha Δ^m podem ser descritos como:

$$\bar{L}^m = T^m + \lambda_m v^m$$

  onde $v^m \doteq T^m - S^m$ é o vector direcção da linha Δ^m.

- Pretendemos encontrar uma posição no espaço correspondente à fonte de luz L, cuja distância quadrática média às rectas Δ^m seja mínima. Define-se o funcional de custo:

$$C(L, \lambda) \doteq \sum_{m=1}^{M} \|\bar{L}^m - L\|^2 \tag{2}$$

  onde $\lambda = [\lambda_1, \ldots, \lambda_M]^T$

A minimização deste funcional tem solução analítica. De notar que para obter L é necessário obter como passo intermédio o vector de parâmetros λ.

6 Organização do Trabalho

Na primeira fase do trabalho de laboratório os alunos terão que resolver os seguintes problemas:

- estimação de pose da câmara com medição manual dos pontos sobre a grelha de calibração;
- estimação da posição da fonte de luz com medição manual dos pontos da sombra do lápis;
- estimação da posição 3D de pontos sobre o objecto com medição manual de pontos do contorno de sombra sobre o objecto.

Na segunda fase do trabalho será efectuada a medição automática dos pontos do contorno de sombra.

São fornecidos conjuntos de imagens associados a cada problema, com os quais os alunos desenvolverão os seus algoritmos. Recomenda-se a utilização do MATLAB [1] para o desenvolvimento dos algoritmos. A meio do semestre, os alunos terão que demonstrar as soluções encontradas para a primeira fase do trabalho. A demonstração terá que focar uma série de sub-objectivos que são descritos em seguida.

6.1 Pose da câmara

- Imagem da grelha de calibração: `patt.bmp`
- A grelha de calibração é constituida por um tabuleiro de xadrez com 64 casas (8x8), cada uma com 3 cm de lado.
- A câmara utilizada tem uma lente com distância focal de 16 mm.
- Cada *pixel* da imagem tem uma dimensão horizontal de 0.025 mm e uma dimensão vertical de 0.0017 mm.
- A câmara encontra-se a cerca de 1m de distância da grelha de calibração.

Sub-objectivos

1. Visualizar a imagem do padrão de calibração com marcações sobrepostas dos pontos da grelha obtidos manualmente.

    - Medir manualmente (p.ex. com o programa `paint`) os N pontos da grelha em coordenadas da imagem e construir uma matriz p de dimensão 2xN:

    $$p = \left[p^1|p^2|\ldots|p^N\right]$$

    em que $p^n = [x^n, y^n]^T$ e x e y são as coordenadas horizontal e vertical da imagem em *pixels*, com origem no canto superior esquerdo da imagem (x cresce para a direita e y para baixo).
    - Ler a imagem da grelha de calibração em MATLAB:
    `patt=bmpread('patt.bmp');`
    - Para a visualização, fazer, p.ex.:
    ```
    image(patt);
    colormap(gray(256));
    hold on;
    plot(p(1,:),p(2,:),'+');
    ```

2. Construir uma matriz P_0, de dimensão 4xN, contendo as coordenadas homogéneas dos pontos da grelha de calibração, expressos no referencial da secretária $\{\mathcal{D}\}$.

    $$P_0 = \left[P_0^1|P_0^2|\ldots|P_0^N\right]$$

3. Estimar a matriz de transformação.

    - Utilizar a função MATLAB fornecida:
    `>> T0 = plcalib(P0, p)`
    em que `P0` e `p` são os vectores obtidos nos pontos anteriores, devidamente ordenados.

4. Confirmar o resultado por visualização da imagem da grelha com as marcações correspondentes à transformação estimada.

    - Obter a matriz P de dimensão 4xN com a descrição dos pontos da grelha no referencial da câmara $\{\mathcal{C}\}$ para T_0 obtido no ponto anterior.
    - Obter a projecção desses pontos no plano de imagem, utilizando a projecção perspectiva.
    - Visualização (idêntica ao ponto 1)

5. Desenvolver um função MATLAB que calcule a intersecção de um raio óptico (O, p)) com o plano da secretária, em coordenadas do referencial da câmara, p.ex.:
`P = plane(p)`

6.2 Posição da fonte de luz

- Imagens de calibração: `lap*.bmp`
- O lápis tem um comprimento de 14.2 cm.

Sub-objectivos

1. Visualizar as imagens de calibração com marcações na base e na extremidade da sombra do lápis.

    - Medir manualmente os pontos na imagem correspondentes à base e à extremidade da sombra do lápis.

- Criar matrizes b (base) e s (sombra da extremidade), de dimensão 2xM contendo os pontos medidos:

$$b = \left[b^1|b^2|\ldots|b^M\right]$$
$$s = \left[s^1|s^2|\ldots|s^M\right]$$

- Para cada imagem sobrepôr as marcações correspondentes.

2. Calcular as matrizes S e T contendo as posições 3D da base e da extremidade do lápis. Projectar os pontos T^m na imagem e confirmar os resultados.

3. Definir o funcional de custo a minimizar em função de $L = [L_x, L_y, L_z]^T$ e de $\lambda = [\lambda_1, \ldots, \lambda_m]$ (ver eq. 2).

4. Minimizar explicitamente esse funcional.

   - Calcular as derivadas parciais do funcional relativamente a L e λ:
   $$\left[\frac{\partial C}{\partial L}, \frac{\partial C}{\partial \lambda}\right]$$
   - Igualar as derivadas parciais a zero e resolver o sistema de equações obtido.

5. Confirmar os resultados obtidos, por visualização – calcule a projecção na imagem da recta que passa por L e por S^m em cada situação m, e sobreponha-a nas imagens de calibração.

6.3 Reconstrução

- Conjuntos de imagens: `cena*.bmp, bulb*.bmp, hor*.bmp`

Sub-objectivos

1. Escolher um dos conjuntos de imagens.

2. Definir o plano de sombra para cada imagem.

   - Medir manualmente dois pontos de sombra sobre a secretária (p_{bot} e p_{top}). Ter cuidado com o facto de a zona de sombra ser bastante larga. Aconselha-se a medição de pontos junto à fronteira da sombra.
   - Calcular a posição 3D (P_{bot} e P_{top}) desses pontos.
   - Definir o plano de sombra em função dos pontos P_{bot}, P_{top} e L, tal que seja fácil calcular a intersecção de um eixo óptico com esse plano (ver anexo A).
   - Desenvolver um função MATLAB que calcule a intersecção de um raio óptico (O, p)) com o plano de sombra, em coordenadas do referencial da câmara, p.ex.:
   `P = shadow(p)`

3. Reconstruir pontos sobre o objecto.

   - Medir manualmente alguns pontos de sombra sobre o objecto, p_k , $k \in [1, \ldots, K]$. Ter cuidado com o facto de a zona de sombra ser bastante larga. Aconselha-se a medição de pontos junto à fronteira da sombra.
   - Obter a posição 3D desses pontos, $P_k = [X_k, Y_k, Z_k]^T$, por intersecção do eixo óptico (O, p_k) com o plano de sombra.

4. Coleccionar os pontos reconstruídos, para todas as imagens, num vector 3xK:
$$P = [P_1|P_2|\ldots|P_K]$$

5. Visualizar a reconstrução obtida, utilizando o comando MATLAB:
`>> mesh`
(consultar o *help* do MATLAB).

6.4 Extracção automática do contorno de sombra – Fase 2

Na segunda fase do trabalho procede-se ao cálculo automático dos pontos no contorno de sombra. Serão aplicadas técnicas simples de processamento de imagem e os alunos terão ao seu dispor sequências de imagens completas com as quais poderão efectuar operações espaciais e temporais e obter reconstruções "densas".

O procedimento será constituido por três passos:

1. Obtenção dos pontos de sombra sobre a secretária, para cada imagem da sequência (para cada instante t), utilizando as linhas de referência auxiliares (ver figura 2). O perfil de nível de brilho ao longo de x nessas linhas apresentará um mínimo pronunciado na zona da sombra. Com base nesse perfil, obtêm-se o pontos pretendidos, com os quais se calcula o plano de sombra no instante t, ou seja: $\pi_s\left(t\right)$.

2. Obtenção dos pontos de sombra sobre o objecto. Neste caso a estratégia anterior não funciona dado o objecto poder apresentar sombras e contornos não devidos ao plano de sombra forçado pelo utilizador. Uma estratégia melhor consiste em:

   - para cada *pixel* da imagem, medir a evolução do nível de brilho ao longo do tempo.

   O perfil medido apresentará um mínimo aquando da passagem do plano de sombra. A partir deste perfil é possível calcular o instante (t_s) de passagem do contorno de sombra, para cada *pixel* (p) da imagem: $t_s\left(p\right)$.

3. Utilizar a informação dos dois pontos anteriores e calcular a posição 3D correspondente a cada *pixel* (p) da imagem, ou seja, intersectar cada raio óptico (O, p) com o plano de sombra correspondente no tempo: $\pi_s\Big(t_s\left(p\right)\Big)$

7 Calendarização

Cada grupo deverá apresentar os resultados da primeira fase do projecto no dia 8 de Maio de 2002.

A apresentação da segunda fase do projecto e a entrega do relatório será efectuada no dia 12 de Junho de 2002.

Note que a avaliação do trabalho terá em conta os resultados da primeira fase, os resultados da segunda fase e o relatório.

8 Recomendações Finais

- Como princípio geral, tente evitar ciclos `for` em MATLAB. Estes ciclos são bastante dispendiosos em termos de tempo de processamento e podem ser evitados quase sempre. Utilize sempre que possível matrizes e operações simples entre elas (o MATLAB está optimizado para operações matriciais).

- Note que as operações envolvidas neste trabalho não são complicadas mas são constituidas por muitos pequenos passos. É portanto recomendado que comece a trabalhar tão cedo quanto possível e retire as suas dúvidas com os docentes assim que as tenha.

A Representação do plano de sombra

- O plano de sombra (π_s) pode ser representado por um vector (K) ortogonal ao plano, e por um escalar (V), definidos por:

$$K \doteq \Big(P_{\text{top}} - L\Big) \times \Big(P_{\text{bot}} - L\Big)$$
$$V \doteq K \cdot L$$

onde $(\times)$ denota o produto externo e $(\cdot)$ o produto interno. O valor absoluto de V equivale ao volume do paralelipípedo $\left\{\overline{OP}_{\mathrm{top}}, \overline{OP}_{\mathrm{bot}}, \overline{OL}\right\}$.

- Qualquer ponto P do plano de sombra π_s verifica a relação:

$$P \cdot K = V \tag{3}$$

Referências

[1] K. Sigmon : MATLAB Primer, University of Florida, 1992.

[2] B. Horn : Robot Vision, MIT Press, McGraw Hill, 1986.

Alexandre Bernardino e José Santos–Victor
Secção de Sistemas e Controlo
Departamento de Engenharia Electrotécnica e de Computadores
Instituto Superior Técnico

Março 2002