UNIVERSIDADE FEDERAL DE MINAS GERAIS — UFMG
INSTITUTO DE CIÊNCIAS EXATAS — ICEx
DEPARTAMENTO DE CIÊNCIA DA COMPUTAÇÃO — DCC

Disciplina:	DCC884 — Visão Computacional
Professor:	Mario Fernando Montenegro Campos (`mario@dcc.ufmg.br`)
Monitor:	Vilar Fiuza da Camara Neto (`neto@dcc.ufmg.br`)
Período:	1º semestre / 2007
Aulas:	Segundas e quartas, das 11:10 às 12:50, na sala 2029 do ICEx
Página:	`http://www.dcc.ufmg.br/verlab/doku.php?id=cursos:visao:index`

4ª Lista de Exercícios

Data da entrega: 28/mai/2007

A lista é individual e as datas de entrega são fixas, devendo o trabalho ser entregue no início da aula. Cada lista vale 5 pontos na nota final. Para cada dia de atraso na entrega será descontado 0,5 ponto do valor da lista.

Esta lista de exercícios é composta de duas partes: a primeira de exercícios teóricos e a segunda com implementações práticas. Não há formato específico para apresentar as respostas e resultados. Porém, clareza na redação é um requisito essencial. Embora o resultado possa ser manuscrito, lembre-se de que "o que não se pode ler não se pode avaliar". Assim, sugere-se que o trabalho seja digitado, podendo o desenvolvimento matemático ser manuscrito.

O Matlab ou Scilab são recomendados para o desenvolvimento das aplicações, pois permitem o desenvolvimento rápido de aplicativos e já incorporam várias rotinas matemáticas e de manipulação de imagens. O Scilab é gratuito e pode ser baixado da página `http://www.scilab.org/`. Você também pode usar outras linguagens de programação, mas nesse caso lembre-se de incluir as instruções de compilação dos programas e eventuais arquivos auxiliares, como "makefiles" ou arquivos de projeto.

Em todos os casos, não se exige nenhuma interface gráfica: os programas podem ler os dados de entrada por argumentos de linha de comando ou por digitação pelo usuário, e podem gravar a saída em arquivos no disco.

Para arquivos gráficos, recomenda-se o uso dos formatos PNG ou JPEG. Nesse sentido o Matlab e o Scilab ajudam, pois ambos disponibilizam rotinas de leitura e gravação para esses formatos. Para os entusiastas de C e C++, a biblioteca FreeImage (`http://freeimage.sourceforge.net/`) é uma alternativa gratuita e multi-plataforma. Para a visualização dos arquivos, os aplicativos IrfanView (Windows) e xv (Linux) são boas ferramentas e também são gratuitos.

Em qualquer situação, os programas devem ser entregues com documentação de uso ou devem ser auto-explicativos.

1 Exercícios Teóricos

1. Descreva e comente sobre os tipos de ruídos que podem afetar a precisão do processo de calibração de câmeras.

2. *Alvo* ou *padrão de calibração* é o nome dado ao objeto ou figura que é imageada para a realização da calibração da câmera. Descreva e comente algumas características desejáveis para um alvo de calibração.

3. De posse dos parâmetros de calibração (intrínsecos e extrínsecos) de uma dada câmera, e dada uma única imagem obtida por essa câmera, é possível descobrir alguma coisa sobre a geometria da cena? O que pode ser feito para recuperar a estrutura tridimensional da cena?

4. Usando o sistema de coordenadas homogêneas, todo o processo de projeção de pontos pode ser representado por uma matriz de transformação M. Assim, a correspondência entre pontos da cena e pontos da imagem pode ser descrita pela relação
$$P_i = M \cdot P_c \ ,$$
sendo P_i as coordenadas em imagem (na forma (x_i, y_i, w_i)) correspondentes à projeção do ponto da cena P_c (na forma (x_c, y_c, z_c, w_c)). Além disso, com coordenadas homogêneas é possível representar pontos no infinito, como $P_i = (x_i, y_i, 0)$ na imagem ou $P_c = (x_c, y_c, z_c, 0)$ na cena.

   Descreva a localização, em relação à câmera, de:

   a) pontos da cena que estão no infinito ($w_c = 0$) que são projetados para pontos de imagem que *não* estão no infinito ($w_i \neq 0$);
   b) pontos da cena que estão no infinito ($w_c = 0$) que são projetados para pontos de imagem que estão no infinito ($w_i = 0$);
   c) pontos da cena que *não* estão no infinito ($w_c \neq 0$) que são projetados para pontos de imagem que estão no infinito ($w_i = 0$).

2 Exercícios Práticos

Se os pixels não são perfeitamente quadrados, então a distância angular entre dois pontos pode cobrir um determinado número de pixels horizontalmente e um diferente número de pixels verticalmente. Nas câmeras CCD padrão, a razão de aspecto (*aspect ratio*) é freqüentemente diferente de uma para outra. Uma forma simples de se determinar essa razão é posicionando um bola de tênis de mesa no centro do campo de visão (*field of view*) e calcular a elongação. Por exemplo, suponha que a imagem da bola possua 30 pixels de largura e 45 pixels de altura. Assim, considerando que a bola seja perfeitamente circular, existem 45 pixels colocados verticalmente no mesmo espaço em que existem 30 pixels posicionados horizontalmente. A razão de aspecto é dada por 45/30 = 1,5, significando que a altura dos pixels é 1,5 vezes maior em relação à largura.

O campo de visão é a distância angular entre dois pontos nas arestas opostas de uma imagem, sejam estas tanto horizontais quanto verticais. Logo, existem dois campos de visão: um vertical e outro horizontal.

Para se medir o campo de visão horizontal, procede-se da seguinte forma: posiciona-se um objeto plano (uma folha de papel com algumas linhas traçadas) em frente à câmera e move-se o objeto até que este preencha, horizontalmente, todo o espaço da câmera. Então mede-se a distância entre a câmera e o objeto como também sua largura. Através de

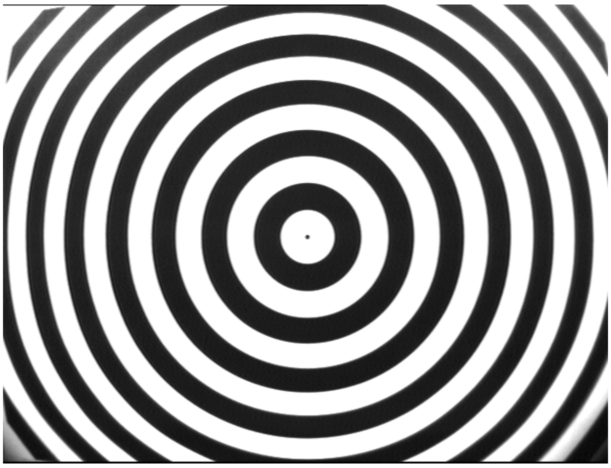

Figura 1: Imagem `circles.png`.

relações trigonométricas simples obtém-se o campo de visão horizontal. Para o cálculo do campo de visão vertical, procede-se de forma semelhante.

Para deixar mais claro o procedimento de cálculo anteriormente descrito, se o objeto possui uma largura d e a distância entre a câmera e o objeto é D, então o campo de visão é dado por

$$\phi = 2 \arctan \left(\frac{d}{2D} \right) . \tag{1}$$

Esta relação fornece os parâmetros iniciais para a calibração da câmera.

Uma folha de papel com círculos concêntricos, posicionada ortogonalmente à direção de visualização da câmera, é imageada de tal maneira que fique quase centralizada em relação ao plano de formação da imagem na câmera. Se a distorção radial é modelada por um polinômio de quarta ordem, então os coeficientes desse polinômio podem ser obtidos através de um procedimento de aproximação linear com base nas distâncias dos centros de cada círculo. Os detalhes que devem ser explorados nessa questão são apresentados a seguir.

Para realizar a calibração, você deve usar a imagem `circles.png` (visto na Figura 1) e algum aplicativo de visualização de imagens para observar a imagem e realizar as medidas sobre ela.

1. A imagem `circles.png` (Figura 1) contém uma imagem produzida em uma folha de papel com anéis concêntricos convenientemente espaçados e imageada em frente à câmera. Use um aplicativo para visualização de imagens para obter as coordenadas dos pixels de interesse, observando com cuidado as convenções da origem do sistema de coordenadas e orientação dos eixos. Essas medidas são suas "medidas de imagem".

   a) Assuma que o centro da imagem é o centro do ponto branco no meio da imagem. Forneça sua melhor estimativa para as coordenadas da linha r e coluna c deste ponto, obtidas através do posicionamento do mouse onde você supõe ser o centro.

   b) Verifique se a razão de aspecto dos pixels é 1 através da medida dos pixels de um círculo e verifique que eles são aproximadamente iguais. Comece com o círculo que você mediu e conhece o centro. Apresente seu resultado e descreva seus cálculos.

c) Determine os campos de visão horizontal e vertical da câmera. Os raios de cada anel medem 1,0 cm e estão espaçados de 1,0 em 1,0 cm. A folha de papel onde os anéis foram impressos foi posicionada a uma distância de 17 cm do centro de projeção da câmera e é ortogonal ao eixo óptico. Descreva seus resultados e seus cálculos.

d) Qual a distância focal, em pixels, usada para essa imagem? Descreva seu raciocínio.

2. Você pode perceber que, nos cantos da imagem, as distâncias entre os círculos maiores parece ser menor do que as distâncias observadas no centro da imagem. Este efeito é causado pela distorção radial assimétrica das lentes da câmera, a qual você deverá determinar. Utilizou-se os coeficientes de distorção para transformar a imagem empregando-se o seguinte modelo:

$$x_d = x\left[1 + (k_0 + k_1\rho^2 + k_2\rho^4)\right] \tag{2}$$

$$y_d = y\left[1 + (k_0 + k_1\rho^2 + k_2\rho^4)\right], \tag{3}$$

onde $\rho = \sqrt{x^2 + y^2}$ é a distância radial ideal em relação ao centro da imagem sem distorção. As coordenadas x_d e y_d são as coordenadas que você mediu na imagem com distorção em relação ao centro. As coordenadas x e y são as coordenadas que deveriam ser medidas sem a presença de distorção. Teoricamente, a constante k_0 pode ser zero. Deixa- se k_0 como um fator para as possíveis correções de erro na medida da distância entre a câmera e o alvo.

a) Para medir as coordenadas horizontais, procede-se uma varredura em uma linha horizontal da imagem, a partir do ponto central, medindo-se as coordenadas relativas as colunas distorcidas c_d, obtendo-se os valores de x_d por $x_d = c_d - c$, onde c representa a coordenada da coluna do ponto central. Escreva o vetor de valores x_d da esquerda para a direita na imagem.

b) Para se obter os valores de x, assume-se que o menor círculo da imagem não está distorcido. Através da interseção das linhas de varredura do meio da imagem temos $x = x_d$. A partir daí, e assumindo que o alvo está posicionado exatamente em frente e paralela em relação a câmera, pode-se inferir outros valores de x. Escreva o resultado em um vetor de valores, varrendo a imagem da esquerda para a direita.

c) Iremos utilizar somente as Equações (2) e (3). Ao longo da linha de varredura média temos $y_d = 0$, portanto a Equação (2) pode ser reescrita como:

$$\delta = x_d - x = k_0 x + k_1 x^3 + k_2 x^5. \tag{4}$$

Assim, uma distorção é levada a um polinômio pela diferença $\delta = x_d - x$. Plote as diferenças δ para os vetores calculados anteriormente.

Determine os coeficientes do polinômio (Equação 4) que melhor aproximam os valores δ que você encontrou anteriormente. Para calcular estes coeficientes,

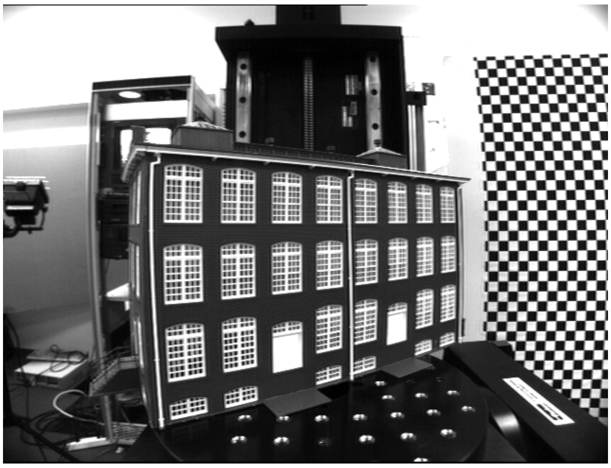

Figura 2: Imagem `lab.png`.

utilize o método dos mínimos quadrados. A idéia da aproximação pelos minímos quadrados consiste em se resolver o sistema de equações lineares $\mathbf{M}\,\mathbf{k} = \delta$.

Seja $\mathbf{x}$ um vetor cujos componentes são todos x, e sejam $\mathbf{x}^3$ e $\mathbf{x}^5$ os dois vetores de terceira e quinta potência com todas as entradas de $\mathbf{x}$. Então, $\mathbf{M} = \left[\mathbf{x}, \mathbf{x}^3, \mathbf{x}^5\right]$. Considere ainda $\mathbf{k}$ como sendo o vetor com as entradas k_0, k_1 e k_2. Os valores δ são os valores anteriormente calculados. Você observará que os valores de k_0 são menores que os valores de k_1, que são ainda menores que os valores de k_2. Isso pode ser evitado se normalizarmos as coordenadas da imagem. Entretanto, se você utilizar estes valores em dupla precisão, não será necessário normalizá-los. (Dica: utilize o Matlab ou outro aplicativo numérico para resolver o sistema).

d) Plote a função δ dada pela Equação (4) na mesma figura que você plotou os valores medidos para δ. Não espere uma correspondência imediata, mas os dois gráficos devem estar muito próximos um do outro.

e) Para testar a calibração, você deve distorcer a imagem `lab.png` (para isso pode-se usar a rotina para Matlab `unwarp.m`, disponível na página da disciplina) e verificar o alinhamento das retas após a operação. A imagem original é apresentada na Figura 2, onde pode-se notar claramente os efeitos da distorção radial: as linhas retas da cena apresentam-se abauladas na imagem. Apresente e comente o resultado obtido.