

Sistema de Realidade Aumentada Baseado em Padrões Rectangulares de Dimensões Desconhecidas

José F. Martins
FEUP, ISMAI

R. Dr. Roberto Frias
4200-465 Porto

jmartins@docentes.ismai.pt

Jorge A. Silva
FEUP, INEB

R. Dr. Roberto Frias
4200-465 Porto

jsilva@fe.up.pt

A. Augusto de Sousa
FEUP, INESC Porto

R. Dr. Roberto Frias
4200-465 Porto

augusto.sousa@fe.up.pt

Sumário

Neste artigo é apresentado um sistema de realidade aumentada baseado no seguimento de padrões rectangulares de dimensões desconhecidas. O conteúdo do padrão a seguir pode ser um qualquer, desde que assimétrico. A partir de uma imagem que contenha o padrão, adquirida no início do processo, é possível calcular a sua razão de forma e gerar vistas frontais do mesmo que são, posteriormente, usadas na fase de seguimento do padrão. Apesar de nos testes efectuados terem sido usados apenas padrões sintetizados, o sistema desenvolvido tem potencial para vir a ser utilizado com padrões naturais.

Palavras-Chave

Realidade Aumentada, Reconhecimento de Padrões, Pontos de Fuga e Visão por Computador.

1. INTRODUÇÃO

A Realidade Aumentada (RA) é um domínio ainda relativamente recente, mas que tem ganho cada vez mais popularidade. Ao longo dos últimos anos surgiram diversas soluções que possibilitam a implementação de ambientes de Realidade Aumentada através de várias técnicas específicas [Bimber05].

Um dos aspectos cruciais de um sistema de RA é a estratégia adoptada para realizar o alinhamento entre o mundo virtual e o mundo real. Actualmente, existem diversas estratégias possíveis para resolver este problema que se baseiam num conjunto diversificado de sensores: o GPS, o giroscópio, a câmara de vídeo, o acelerómetro e outros [Rolland01]. As soluções mais comuns baseiam-se no reconhecimento de um padrão colocado na cena, numa sequência de imagens de vídeo, recorrendo a técnicas de Visão por Computador. Este padrão serve de base ao processo de alinhamento entre os objectos reais e virtuais existentes nos respectivos mundos.

Os padrões são vulgarmente sintetizados, ou seja, padrões artificiais propositadamente concebidos com o intuito de facilitar o processo de reconhecimento e estrategicamente colocados no mundo real. No entanto, existem situações em que a utilização deste tipo de padrões se torna muito difícil ou mesmo impossível, nomeadamente em ambientes exteriores. Por exemplo, no caso de se pretender colocar uma chaminé virtual sobre um telhado de um edifício real, seria impraticável colocar um destes padrões no telhado do edifício, não só pela óbvia dificuldade de acesso, como também pela dimensão do mesmo.

Nestas situações, a utilização de padrões naturais (não sintetizados) que existam no mundo real e portanto não propositadamente criados para o efeito, parece ser uma opção mais adequada. No exemplo anterior, um padrão natural poderia ser definido com base numa janela ou em outro detalhe do edifício que possa sobressair visualmente e que possa ser reconhecido como um padrão. No entanto, o processo de reconhecimento destes padrões é geralmente bastante mais complexo, não sendo tão facilmente identificados como os padrões sintetizados.

O processo de alinhamento recorrendo a padrões naturais é igualmente complexo. Este processo necessita que seja estabelecida a correspondência entre um conjunto de pontos existentes no padrão e as respectivas projecções existentes na imagem adquirida. Para que seja estabelecida a correspondência, é importante conhecer a forma geométrica do padrão e as respectivas dimensões reais ou, pelo menos, a sua razão de forma.

Os padrões sintetizados possuem formas e dimensões conhecidas, assumindo frequentemente uma forma quadrangular. Por outro lado, os padrões naturais possuem potencialmente uma grande diversidade de formas, muitas delas irregulares; além disso, frequentemente, as respectivas dimensões são desconhecidas.

No presente trabalho, com o intuito de limitar a diversidade de formas a serem processadas, só são considerados padrões com formas rectangulares. Realmente, o rectângulo é uma forma geométrica muito comum no mundo real, sendo particularmente abundante em edifícios e noutras estruturas criadas pelo ser humano. Na maioria das situações, é difícil obter as dimensões reais destes rectângulos mas, em contrapartida, as respectivas razões

de forma podem ser obtidas sem que as dimensões reais sejam conhecidas.

Para determinar a razão de forma de um rectângulo, basta conhecer a relação de proporção entre as suas dimensões. A principal dificuldade advém da forma geométrica que o rectângulo real apresenta após a sua projecção no plano de imagem. Devido às distorções causadas pela perspectiva, não é possível inferir directamente a razão de forma do rectângulo a partir das dimensões do quadrilátero resultante de uma projecção na imagem.

De acordo com [Alvarez02], é possível recuperar a razão de forma de um objecto rectangular de dimensões desconhecidas, mesmo em situações em que o eixo óptico da câmara não esteja orientado perpendicularmente em relação à superfície onde assenta o padrão. Estas situações, muito habituais na captura de padrões naturais, podem ser causadas por limitações físicas impostas por obstáculos ou pela própria configuração do espaço envolvente ao padrão.

Neste artigo propõe-se uma solução para o problema do alinhamento através de padrões rectangulares de dimensões desconhecidas. O sistema de RA desenvolvido é capaz de calcular a razão de forma de um padrão rectangular desconhecido a partir da sua projecção numa imagem adquirida no início do processo. Apesar deste sistema utilizar unicamente padrões sintetizados, prevê-se que no futuro possa vir a utilizar padrões naturais.

Para além da introdução, este documento é composto pelas seguintes secções: o Estado da Arte, na qual se faz um resumo de alguns sistemas de RA existentes, com particular incidência nos que efectuem o processo de alinhamento sem o recurso a padrões sintetizados; o Sistema Proposto, onde é descrito o sistema de RA e o cálculo da razão de forma de um padrão rectangular projectado no plano de imagem a partir da determinação dos pontos de fuga dessa imagem; a Implementação que descreve o desenvolvimento do sistema de RA; a Avaliação, onde são apresentados os resultados obtidos e, finalmente, as Considerações Finais, na qual são retiradas as principais considerações sobre o trabalho realizado.

2. ESTADO DA ARTE

Um dos primeiros sistemas de RA e um dos mais conhecidos e difundidos, na actualidade, é o ARTToolkit [ARTToolkit07] que surgiu inicialmente como uma componente de um sistema de videoconferência em RA [Kato99].

O ARTToolkit é um sistema de RA que realiza o alinhamento através de padrões sintetizados, de forma quadrada, constituídos por uma borda exterior negra que delimita uma representação visual assimétrica de elevado contraste.

Apesar de ser incapaz de lidar com padrões naturais, este sistema continua a ser uma importante referência no domínio da RA.

No entanto, durante a presente década, surgiram algumas soluções que realizam o processo de alinhamento através

de técnicas que não envolvem padrões artificiais. Destes sistemas salientam-se os seguintes.

O sistema de RA apresentado por Ferrari [Ferrari01] é baseado em padrões naturais que consistem em rectângulos ou elipses pertencentes a superfícies planas no mundo real. Estes padrões são inicialmente identificados pelo utilizador e devem corresponder a zonas de um plano real que possuam detalhes visuais que facilitem a sua detecção.

O sistema de RA de Simon [Simon02] é baseado no seguimento de pontos característicos existentes num plano identificado inicialmente pelo utilizador. Neste sistema, o utilizador deve identificar manualmente quatro ou mais pontos na imagem que correspondem a pontos coplanares no mundo real, estes pontos vão delimitar uma área na qual o sistema vai procurar pontos característicos que sejam facilmente detectados ao longo do vídeo capturado. O utilizador deve igualmente alinhar manualmente o sistema de referência virtual com o plano real identificado anteriormente, através da inserção de um objecto virtual na imagem adquirida, tem ainda a liberdade de transformar o objecto virtual até que o alinhamento do mesmo com o mundo real seja o pretendido.

Em [Comport03] é apresentado um sistema de RA baseado no reconhecimento de linhas, círculos, cilindros ou esferas existentes no mundo real. As linhas podem identificar as arestas de objectos 3D, como um cubo ou objectos mais complexos, como por exemplo: uma cadeira. No entanto, este sistema necessita que o utilizador, numa fase inicial, identifique manualmente as entidades reais existentes na imagem que servirão como mecanismo de alinhamento.

O sistema de RA de Frahm [Frahm05] permite o alinhamento de objectos virtuais numa cena real sem recorrer a qualquer tipo de padrão. A posição e orientação da câmara são estimadas a partir do seguimento de pontos característicos existentes nas imagens adquiridas. Este sistema usa uma segunda câmara que é utilizada para a determinação da iluminação existente no espaço real. Dessa forma, os modelos virtuais são iluminados realisticamente e projectam sombras igualmente realistas.

3. SISTEMA PROPOSTO

Descreve-se a seguir o sistema de RA implementado que permite fazer o alinhamento com base em padrões sintetizados, rectangulares de conteúdo e dimensões desconhecidos.

O sistema é constituído por três componentes fundamentais: a Calibração da Câmara; a Extracção e Identificação do Padrão e a Geração de Imagens Aumentadas.

3.1 Calibração da Câmara

A primeira componente do sistema é responsável pelo cálculo dos parâmetros intrínsecos da câmara e dos respectivos coeficientes de distorção. Os parâmetros intrínsecos são importantes para o cálculo da posição e da orientação da câmara que é realizado na componente final. Os coeficientes de distorção possibilitam a correc-

ção da imagem adquirida, removendo a distorção criada pelas lentes da câmara. Doravante, esta correcção será efectuada em todas as imagens adquiridas.

O processo de calibração da câmara, que será descrito com mais detalhe na secção 4.1, é relativamente moroso e geralmente é efectuado uma vez para cada câmara (assumindo que os parâmetros intrínsecos da mesma se mantêm inalterados).

3.2 Extração e Identificação do Padrão

O sistema proposto utiliza padrões sintetizados rectangulares. Estes padrões são constituídos por uma borda negra que delimita uma zona interior que é constituída por uma determinada representação visual assimétrica.

Nesta componente, o sistema de RA calcula a razão de forma do padrão e efectua a captura da zona interior do mesmo.

Para efectuar o alinhamento é necessário estabelecer a correspondência entre os cantos do padrão no mundo real e as respectivas projecções na imagem adquirida.

As coordenadas dos cantos do padrão na imagem são obtidas através de algoritmos de processamento e análise de imagem. Estas coordenadas são expressas em *pixels* relativamente ao sistema de referência da imagem.

Por outro lado, as coordenadas dos cantos do padrão no mundo real são relativas ao sistema de referência do padrão. A origem deste sistema assume-se como estando posicionada no centro geométrico do padrão, sendo o respectivo eixo dos z perpendicular ao plano em que o padrão se encontra; os eixos x e y , por uma questão de simplificação, são paralelos aos lados do padrão. A unidade utilizada pode ser uma unidade real (milímetros, centímetros, polegadas, etc.) ou pode, em certos casos, ser uma unidade genérica relativa a uma das dimensões do padrão: a sua altura ou a sua largura.

A principal vantagem da utilização da unidade genérica é que esta não necessita das dimensões reais do padrão mas, ainda assim, necessita de conhecer a sua razão de forma. Por exemplo, no caso de um rectângulo com uma razão de 4:3, se a sua altura corresponder à unidade genérica então a sua largura iria corresponder aproximadamente a 1.333 unidades genéricas.

A razão de forma é, assim, imprescindível para o cálculo das coordenadas dos cantos do padrão no mundo real relativas ao sistema de referência do padrão. Sem estas coordenadas, não será possível estabelecer a correspondência entre os cantos do padrão no mundo real e as respectivas projecções na imagem adquirida que é a base do processo de alinhamento.

Adicionalmente, a razão de forma é igualmente importante para a captura da zona interior do padrão existente na imagem adquirida. Um dos problemas que se põe nos sistemas de RA baseados no seguimento de padrões é a determinação da orientação do padrão. Para determinar essa orientação, é necessário comparar a zona interior do padrão existente na imagem adquirida com quatro imagens armazenadas previamente, em pose frontal, com

orientações diferentes. Para isso, a zona interior deve ser assimétrica, para que não existam ambiguidades na determinação dessa orientação.

Para obter a zona interior do padrão, em pose frontal, é necessário conhecer a respectiva razão de forma.

3.2.1 Recuperação da Razão de Forma

O método utilizado para recuperar a razão de forma do padrão, a partir de uma única imagem do mesmo, ainda que o padrão surja distorcido pela perspectiva, baseia-se na identificação dos pontos de fuga da imagem [Cipolla99]. O método já foi utilizado com sucesso noutros contextos, nomeadamente para a inserção de objectos 3D em fotografias de edifícios [Alvarez02] e para a modelação 3D baseada em imagem [P3D07].

Inicialmente, há que identificar pares de rectas na imagem adquirida, para cada um dos pontos de fuga, que correspondam a rectas paralelas ou perpendiculares ao plano horizontal do mundo real. Cada um dos pontos de fuga (F_x , F_y e F_z) resulta da intersecção de um par de rectas e as respectivas coordenadas são expressas em *pixels*, no sistema de referência da imagem. Por sua vez, a identificação das rectas pode ser realizada, por uma questão de simplificação, através da identificação de seis pontos (V_0 a V_5) que coincidem com os vértices que delimitam as três arestas verticais visíveis de um paralelepípedo. Por exemplo; na figura 1, o ponto de fuga F_x resulta da intersecção da recta que passa pelos vértices V_0 e V_1 com a recta que passa pelos vértices V_2 e V_3 . Os restantes pontos de fuga são calculados de forma semelhante.

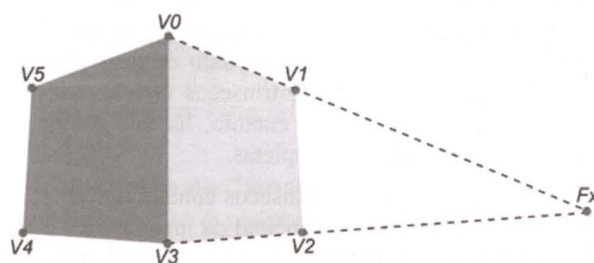


Figura 1: Cálculo do ponto de fuga F_x através da intersecção da recta V_0V_1 com a recta V_2V_3 .

De acordo com [Cipolla99], a posição do centro óptico (C) pode ser calculada a partir dos pontos de fuga anteriores. O centro óptico é o resultado da intersecção de três rectas que passam respectivamente por cada um dos pontos de fuga e que formam um ângulo recto entre si no ponto de intersecção (ver figura 2). As coordenadas do centro óptico (c_x , c_y e c_w) são expressas em *pixels*, no sistema de referência da imagem, e a coordenada w que representa a distância deste ponto ao plano da imagem é a distância focal. Adicionalmente, a projecção do centro óptico no plano de imagem (C') é o centro ortogonal de um triângulo formado pelos três pontos de fuga anteriormente calculados.

Em seguida, são criados três vectores (\vec{CF}_x , \vec{CF}_y e \vec{CF}_z) que ligam o centro óptico a cada um dos pontos de fuga. Estes vectores são perpendiculares entre si e constituem,

após a normalização, um sistema de referência com a origem no centro óptico. Este é considerado o sistema de referência do mundo e é exclusivamente utilizado na recuperação da razão de forma do padrão rectangular, não sendo utilizado na geração de imagens aumentadas.

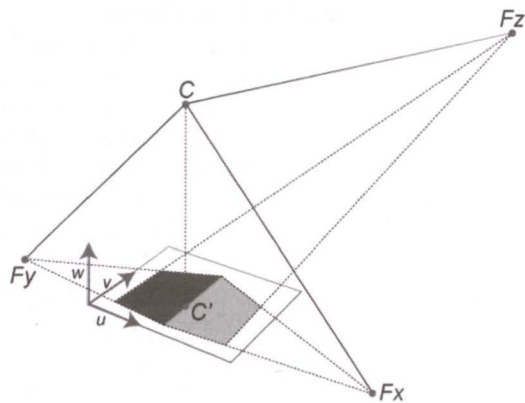


Figura 2: Cálculo do centro óptico (C) a partir dos pontos de fuga.

O sistema de referência definido tem a particularidade dos respectivos planos ordenados serem paralelos às faces de um paralelepípedo utilizado para o traçado dos pares de recta que conduziram aos pontos de fuga. Os planos YZ e XZ são paralelos às faces verticais do paralelepípedo e o plano XY é paralelo à base e ao topo (faces horizontais) do paralelepípedo.

A projecção de um ponto do sistema de referência do mundo real (X,Y,Z) no sistema de referência da imagem (u,v) é realizada através da equação de projecção da câmara (1). Esta equação é o produto de duas matrizes: uma matriz de parâmetros intrínsecos e uma matriz de parâmetros extrínsecos. No entanto, há que referir que ambas as matrizes são incompletas.

A matriz de parâmetros intrínsecos contém unicamente a distância focal e o ponto principal da imagem, e a matriz de parâmetros extrínsecos resume-se a uma matriz de rotação. A justificação para a não inclusão de outros parâmetros intrínsecos, como por exemplo, os coeficientes de distorção, é explicada pelo facto de esta abordagem simular uma câmara do tipo *pin-hole*, na qual não é considerada a existência de lentes. Isto implica que os coeficientes de distorção tenham já sido calculados e que a imagem adquirida utilizada para a identificação das linhas de fuga tenha sido previamente corrigida no sentido de cancelar os efeitos da distorção das lentes.

Relativamente à matriz de parâmetros extrínsecos, a não inclusão do vector de translação é justificada pelo facto de a origem do sistema de referência do mundo ser coincidente com o centro óptico da câmara.

O ponto principal da imagem consiste na projecção do centro óptico no plano de imagem (c_u e c_v); a distância focal calcula-se pela distância do centro óptico ao plano de imagem (c_w). Por sua vez, a matriz de parâmetros extrínsecos é constituída pela matriz de rotação que con-

siste nas componentes dos vectores $CF_x=(u_1,v_1,w_1)$, $CF_y=(u_2,v_2,w_2)$ e $CF_z=(u_3,v_3,w_3)$ dispostos segundo as colunas da matriz.

$$\begin{bmatrix} tu \\ tv \\ t \end{bmatrix} = \begin{bmatrix} C_w & 0 & C_u \\ 0 & C_w & C_v \\ 0 & 0 & I \end{bmatrix} \begin{bmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \quad (1)$$

À partida, não é possível obter as coordenadas dos cantos do padrão no mundo real a partir das respectivas projecções na imagem adquirida, uma vez que existe um número infinito de pontos 3D cujas projecções são coincidentes com um mesmo ponto 2D na imagem. Torna-se no entanto possível, se uma das coordenadas 3D de cada canto for conhecida.

O processo passa por, em primeiro lugar, identificar a face do paralelepípedo na qual o padrão se encontra. Ao identificar esta face é possível identificar qual o plano ordenado que lhe é paralelo. Se, por exemplo, o plano do padrão for paralelo ao plano ordenado XZ então poderemos afirmar que as coordenadas y de todos os pontos existentes no plano do padrão são iguais. Um procedimento equivalente pode ser aplicado às restantes faces do paralelepípedo.

Visto que, no sistema referencial anterior, as coordenadas y dos cantos do padrão são iguais, pode-se, por simplificação, e levando em consideração que não se pretende obter as reais dimensões do padrão, mas sim, a sua razão de forma, assumir que os valores das coordenadas y são iguais à unidade. Assim sendo, o problema de encontrar um único ponto 3D que corresponda ao ponto 2D projectado no plano de imagem é consideravelmente simplificado, uma vez que uma das coordenadas do ponto 3D é conhecida (no exemplo dado, é a coordenada y).

É então possível, através da inversa da matriz de projecção da câmara, obter as coordenadas dos pontos 3D do mundo real a partir das respectivas projecções existentes na imagem adquirida. Com esta informação pode, em primeiro lugar, verificar-se se a forma geométrica obtida é efectivamente um rectângulo e, em caso afirmativo, efectuar o cálculo da respectiva razão de forma.

3.2.2 Extração do Padrão

Uma vez que o padrão a seguir não é suposto ser conhecido à priori, torna-se necessário que, numa fase inicial do funcionamento do sistema, se capture uma imagem que possa vir a ser usada como padrão.

Na impossibilidade de a captura ser feita em pose frontal, é necessário que a imagem capturada seja processada de forma a eliminar a distorção da perspectiva. Designa-se esta fase por "Extração do Padrão".

Esta fase começa por isolar a zona interior do padrão numa imagem adquirida, identificando os cantos interiores do bordo negro que o delimita; obtém-se assim um rectângulo distorcido pelo efeito de perspectiva. Com base nas coordenadas dos cantos da zona interior do

padrão na imagem adquirida e na razão de forma anteriormente calculada, é possível determinar os parâmetros de uma matriz de homografia¹ que permite transformar a zona interior distorcida numa zona interior em pose frontal (figura 3).

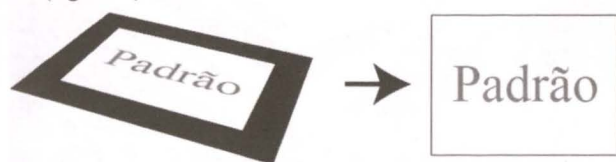


Figura 3: Transformação da zona interior do padrão numa zona interior em pose frontal.

Posteriormente é memorizada uma cópia da zona interior que servirá como o padrão do sistema. De forma, a melhorar a eficiência do sistema, são memorizadas quatro cópias, com orientações diferentes (0°, 90°, 180° e 270°).

3.3 Geração de Imagens Aumentadas

A terceira e última componente é responsável pela fase activa da RA. Nesta fase, os objectos virtuais são alinhados com os objectos reais através do padrão existente na cena real. Portanto, o primeiro passo consiste em realizar a detecção do padrão na imagem adquirida.

Levando em consideração que os padrões a serem detectados são padrões sintetizados, o processo de detecção será consideravelmente facilitado relativamente a um processo equivalente relativo a padrões naturais. Este processo será explicado, em detalhe, na secção de Implementação.

A detecção do padrão vai possibilitar a identificação das coordenadas dos respectivos cantos e a captura da zona interior do mesmo. As coordenadas dos cantos do padrão na imagem adquirida conjuntamente com as coordenadas desses mesmos cantos no sistema de referência do padrão vão ser utilizados no cálculo de uma matriz de homografia que vai possibilitar o cálculo dos parâmetros extrínsecos da câmara.

Para calcular esta matriz de homografia é necessário determinar a orientação do padrão relativamente à câmara, para garantir que a correspondência entre os cantos do padrão na imagem com os cantos do mesmo no sistema de referência do padrão seja a correcta. Para tal, é necessário comparar a zona interior do padrão existente na imagem adquirida com as quatro imagens criadas na componente anterior. Convém salientar que estas quatro imagens possuem duas razões de forma inversas, as imagens com uma rotação de 0° e 180° possuem uma razão igual à razão calculada na etapa anterior, enquanto que as restantes imagens possuem uma razão inversa desta última. A zona interior do padrão existente em cada imagem adquirida deve ser transformada em duas imagens retangulares, uma com a razão igual à calculada anteriormente e a outra com uma razão inversa desta última.

¹ Esta matriz permite realizar o mapeamento entre pontos de duas imagens de um mesmo plano 3D, obtidas com duas câmaras em posições diferentes.

Cada uma destas imagens é comparada com as imagens análogas, armazenadas na componente anterior.

Para realizar a síntese gráfica dos objectos virtuais de forma a apresentarem um alinhamento visual com o padrão existente no mundo real é necessário criar uma câmara virtual que possua os mesmos parâmetros intrínsecos e extrínsecos da câmara real. Os parâmetros intrínsecos foram calculados na componente de calibração da câmara e os parâmetros extrínsecos são calculados, para cada imagem adquirida, na componente de geração de imagens aumentadas. Conhecidos ambos os parâmetros, é possível determinar a matriz de projecção de perspectiva da câmara.

4. IMPLEMENTAÇÃO

O sistema de RA apresentado foi implementado, em grande medida, recorrendo a funcionalidades existentes na API OpenCV [OCV07] e na API OpenGL [OGL07] bastante utilizadas no desenvolvimento de aplicações nos domínios da Visão por Computador e da Computação Gráfica.

O sistema de RA encontra-se dividido em três componentes fundamentais: a componente de calibração da câmara; a componente da extracção e identificação do padrão e a componente de geração de imagens aumentadas. As duas primeiras componentes fazem uso de funcionalidades² do OpenCV e correspondem à fase inactiva da RA. A última componente faz uso de funcionalidades de ambas as API's e corresponde à fase activa da RA.

Este sistema foi desenvolvido e testado num PC com um processador *Intel Pentium 4* de 2.4 GHz, com 2.5 GB de memória RAM, com uma placa gráfica *MSI nVidia GeForce4 Ti4800-SE8X* com 128 MB de memória RAM e suporte de OpenGL 1.3. A *webcam* utilizada foi uma *Creative Live! Cam Vista IM* com uma resolução máxima de 800x600.

4.1 Calibração da Câmara

Nesta componente é realizada a calibração da câmara. Para isso recorre-se à *cvCalibrateCamera2*. Esta funcionalidade realiza o cálculo dos parâmetros intrínsecos da câmara.

Para isso é feita a captura de uma série de vistas distintas (de diferentes orientações e distâncias) de uma grelha de calibração, que consiste numa folha de papel A4 com um padrão de xadrez composto por 9x7 quadrículas de lado igual a 27 mm (figura 4).

² Todas as expressões, em itálico, que comecem por "cv" consistem em funcionalidades da API OpenCV.

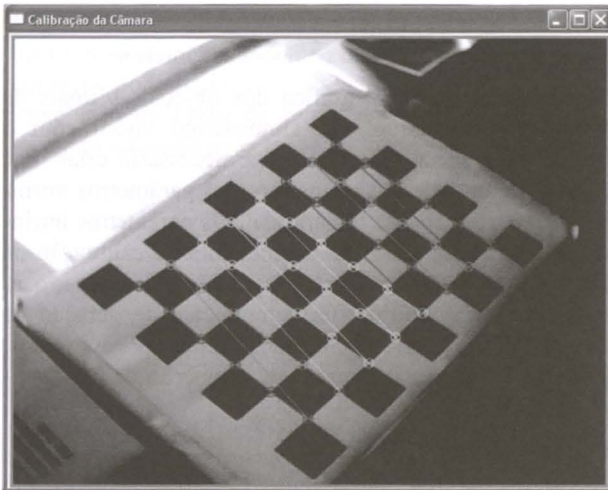


Figura 4: Calibração da câmara recorrendo à respectiva grelha.

A detecção dos pontos de calibração (cantos das quadrículas) é realizada automaticamente através da *cvFindChessboardCorners*. Após a repetição deste processo, para as diversas vistas da grelha de calibração, é então possível realizar o cálculo dos parâmetros intrínsecos através da *cvCalibrateCamera2*. Convém salientar que os coeficientes de distorção da câmara (radiais e tangenciais) são igualmente calculados através desta função. Estes coeficientes são utilizados nas duas componentes seguintes com o intuito de corrigir a distorção induzida pela lente da câmara na imagem adquirida. Esta correcção é realizada recorrendo à *cvUndistort2*.

Os parâmetros intrínsecos são armazenados num ficheiro, possibilitando o seu uso futuro, sem necessidade de recalibração, desde que não sejam feitos ajustes na câmara que impliquem alterações dos mesmos (por exemplo, o ajuste da distância focal).

4.2 Extração e Identificação do Padrão

Esta componente é responsável pelo cálculo da razão de forma do padrão e pela captura da respectiva zona interior (a zona delimitada pela borda exterior).

Este cálculo exige, numa primeira fase, que o utilizador realize a identificação manual de seis vértices de um paralelepípedo de dimensões desconhecidas, estes seis vértices vão definir um conjunto de seis rectas que são utilizadas para identificar os três pontos de fuga que servem de base ao processo descrito na secção 3.2.

Na figura 5 é possível visualizar a identificação dos seis vértices. Ainda nessa figura, o padrão assente no topo do paralelepípedo possui a forma geométrica de um rectângulo com a razão de forma de 4:3 e a zona interior do mesmo é constituída por um calendário de bolso lá colocado.

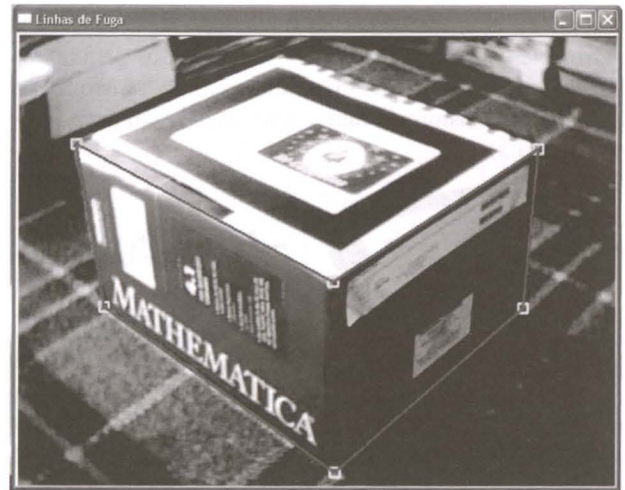


Figura 5: Identificação dos seis vértices do paralelepípedo.

Na eventualidade da inexistência de um paralelepípedo, esta componente possibilita a identificação de doze pontos que vão formar as seis rectas necessárias. Estes pontos necessitam de pertencer a rectas imaginárias que no mundo real sejam paralelas ou perpendiculares ao plano horizontal. Importa salientar que a imagem adquirida, na qual é realizada a identificação dos vértices, foi corrigida das distorções causadas pelas lentes.

Em seguida, o utilizador necessita de identificar qual o plano ordenado que é paralelo ao plano no qual se encontra o padrão (planos: XY , XZ e YZ). É então calculada a razão de forma do padrão e, adicionalmente, é capturada a zona interior do mesmo. Esta captura poderá ser realizada numa outra vista e não necessariamente na mesma que foi utilizada para o cálculo anterior.

Recorrendo a *cvWarpPerspectiveQMatrix* é criada uma matriz de homografia que é utilizada pela *cvWarpPerspective* para transformar a zona interior do padrão distorcido na imagem adquirida numa imagem rectangular com a razão de forma calculada anteriormente e sem as distorções causadas pela perspectiva. Em seguida, são criadas quatro cópias desta imagem com diferentes orientações (ver figura 6).

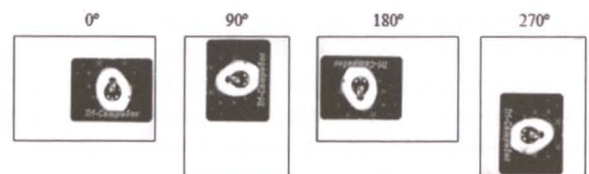


Figura 6: As quatro cópias do padrão capturado, cada uma com a sua respectiva orientação.

Finalmente, é guardada em ficheiro diversa informação relativa ao padrão capturado, nomeadamente a respectiva razão de forma, a relação de proporção entre o contorno exterior da borda e o respectivo contorno interior e as quatro imagens do padrão em pose frontal.

4.3 Geração de Imagens Aumentadas

Esta componente é responsável pela fase activa da RA. Dessa forma, é capaz de reconhecer um padrão na ima-

gem adquirida, calcular os parâmetros extrínsecos da câmara e sintetizar os modelos virtuais devidamente alinhados com o padrão existente na imagem adquirida do mundo real (ver figura 7).

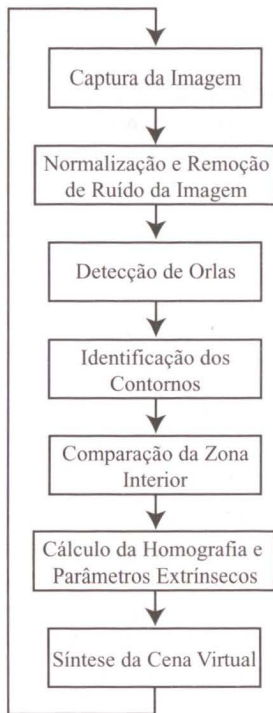


Figura 7: Fluxograma simplificado do processo de geração de imagem aumentadas.

A fase activa da RA implica que as imagens sejam processadas em tempo real. A primeira fase deste processamento tem como objectivo detectar os contornos, interno e externo, da borda do padrão. Estes contornos possuem algumas características importantes: são ambos quadriláteros convexos; o contorno externo contém o interno e existe uma relação de proporção entre as áreas delimitadas por cada um dos contornos.

Inicialmente, a imagem de cor adquirida é convertida para níveis de cinzento. Em seguida, a imagem é normalizada, com o intuito de tornar o algoritmo de detecção do padrão parcialmente invariante ao nível geral de iluminação.

O processo de normalização começa por identificar o tom de cinzento mais escuro e o mais claro existente na imagem. Seguidamente, o intervalo existente entre estes dois limites é ajustado de forma a aproveitar toda a gama de cinzentos, do preto ao branco.

Após o processo de normalização, o nível de ruído existente na imagem é reduzido pela aplicação de um filtro de mediana, de dimensão 3x3, recorrendo a *cvSmooth*.

O primeiro passo, para a detecção dos contornos internos/externos da borda do padrão, consiste na detecção das orlas existentes na imagem. Esta detecção é realizada através do algoritmo de detecção de orlas de Canny [Canny86] disponibilizado pela *cvCanny*.

A partir das orlas detectadas são identificados os contornos fechados existentes na imagem através da *cvFind-*

Contours. Estes contornos são geralmente constituídos por um número elevado de vértices, muitos deles colineares e, dessa forma, redundantes. Para simplificar estes contornos e reduzir o número de vértices ao número de cantos do padrão é utilizada a *cvApproxPoly*.

Do universo de contornos resultantes são filtrados todos os contornos cujo o número de vértices seja diferente de quatro. Dos contornos restantes, são filtrados todos os contornos que não sejam convexos. Finalmente, são filtrados todos os contornos que delimitem uma área demasiado pequena.

O contorno que delimita a maior área é identificado e, caso este último contenha um outro contorno, cuja a relação de proporção entre as respectivas áreas delimitadas seja a correcta para o padrão em causa, então o primeiro contorno é considerado o contorno externo da borda do padrão e o segundo é o contorno interno da mesma. Caso contrário, o contorno com a maior área é considerado o contorno interno da borda.

Identificado o contorno interno, a informação visual existente na zona interior do padrão é transformada em duas imagens rectangulares (uma delas com uma razão de forma igual à calculada anteriormente e a outra com uma razão inversa desta última). Estas duas imagens rectangulares são então comparadas com as quatro cópias da zona interior do padrão armazenadas anteriormente. Cada uma destas imagens rectangulares será unicamente comparada com as duas cópias com a mesma razão. Esta comparação é realizada através da *cvMatchTemplate*. Com esta comparação é possível identificar cada um dos cantos detectados anteriormente.

Se o contorno externo da borda for encontrado, é calculada uma matriz de homografia que relaciona os seus vértices com os cantos do padrão no sistema de referência do próprio. Caso não seja encontrado, os vértices do contorno interior são utilizados. Neste caso, a matriz de homografia vai relacionar estes últimos vértices com os respectivos vértices do contorno interior da borda no sistema de referência do padrão.

A matriz de homografia, assim calculada, serve de base ao cálculo dos parâmetros extrínsecos da câmara para a actual imagem adquirida. Este cálculo é realizado através da *cvFindExtrinsicCameraParams2*. Com os parâmetros intrínsecos e extrínsecos da câmara conhecidos é possível criar virtualmente a câmara através da *glFrustum* e da *glLoadMatrixd*. É, finalmente, possível sintetizar objectos virtuais alinhados com o padrão existente no mundo real (ver figura 8).

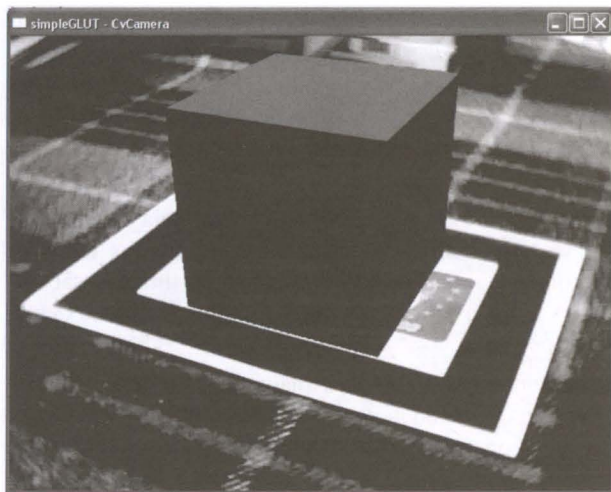


Figura 8: Alinhamento de um cubo virtual com o padrão criado anteriormente.

5. AVALIAÇÃO

A avaliação do sistema de RA desenvolvido incidiu fundamentalmente na recuperação da razão de forma do padrão rectangular e no desempenho do sistema durante a fase activa da RA.

A determinação correcta da razão de forma é um processo importante para o funcionamento do sistema de RA. A avaliação efectuada demonstrou que a resolução da imagem adquirida e o ângulo entre o eixo óptico da câmara e o plano do padrão são os factores que mais influenciam a exactidão do resultado.

Para imagens adquiridas com uma resolução de 640x480 e para ângulos inferiores ou iguais a 45°, a razão de forma é calculada com um erro relativo inferior a 2%. Para ângulos superiores a 45°, o erro relativo aumenta progressivamente (ver tabela 1).

Ângulo	Erro Relativo Médio
5°	0,23%
15°	0,44%
25°	0,78%
35°	1,15%
45°	1,73%
55°	5,35%
65°	13,43%
75°	34,01%

Tabela 1: Erro relativo médio na obtenção da razão de forma.

Durante a fase activa da RA, o sistema é capaz de realizar o alinhamento de objectos virtuais através de padrões sintetizados de diversas formas rectangulares.

Foram utilizados padrões que possuíam razões de forma entre 1:1 (quadrado) e 3:1. Com estes padrões, o sistema é capaz de manter um alinhamento satisfatório, ou seja o desalinhamento é visualmente imperceptível. Para padrões com razões superiores e recorrendo a imagens adquiridas com uma resolução 320x240, a qualidade do alinhamento começa progressivamente a degradar-se.

Usando o sistema de computação anteriormente descrito foi possível processar imagens com uma frequência igual a 25 Hz e adquiridas com uma resolução de 320x240. Para uma resolução de 640x480, o sistema processa as imagens adquiridas com uma frequência na ordem dos 10 Hz-15 Hz.

6. CONSIDERAÇÕES FINAIS

O sistema de RA desenvolvido é capaz de realizar o alinhamento através de padrões rectangulares com dimensões previamente desconhecidas. Esta situação é muito comum num ambiente real, no qual a forma rectangular é particularmente abundante e no qual as respectivas dimensões reais são frequentemente desconhecidas.

A característica mais importante deste sistema é a recuperação da razão de forma do padrão rectangular a partir de uma só imagem adquirida. Para determinadas condições, esta recuperação é realizada com uma exactidão aceitável. A razão assim obtida é fundamental para o processo de alinhamento subsequente.

Actualmente, o sistema funciona unicamente com padrões sintetizados, nos quais a razão de forma poderá ser facilmente determinada. No entanto, o futuro objectivo deste sistema é a utilização de padrões naturais de forma rectangular, na qual a respectiva razão não seja conhecida de antemão.

7. REFERÊNCIAS

- [Alvarez02] B. S. Alvarez, P. C. P. Carvalho e M. Gattass. Insertion of Three-dimensional Objects in Architectural Photos. *Journal of WSCG*, vol. 10, n. 1-3; 10th International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision, 2002.
- [ARToolkit07] Augmented Reality Toolkit Website, Junho de 2007.
<<http://www.hitl.washington.edu/artoolkit>>
- [Bimber05] O. Bimber and R. Raskar. Modern Approaches to Augmented Reality. *International Conference on Computer Graphics and Interactive Techniques*, ACM SIGGRAPH 2005.
- [Camp03] A. Comport, E. Marchand e F. Chaumette. A Real-Time Tracker for Markerless Augmented Reality. *ACM/IEEE Int. Symp. on Mixed and Augmented Reality, ISMAR'03*, Tokyo, Japan, 2003.
- [Canny86] J. Canny. A Computational Approach to Edge Detection. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 1986.
- [Cipolla99] R. Cipolla, T. Drummond e D. Robertson. Camera Calibration from Vanishing Points in Images of Architectural Scenes. *BMVC*, 1999.
- [Ferrari01] V. Ferrari, T. Tuytelaars e L. Van Gool. Markerless Augmented Reality with a Real-Time Affine Region Tracker. *Proc. IEEE and ACM Intl. Symposium on Augmented Reality, volume I*, 2001.
- [Frahm05] J-M. Frahm, K. Köser, D. Grest e R. Koch. Markerless Augmented Reality with Light Source Esti-

mation for Direct Illumination. *Conference on Visual Media Production*, London, 2005.

[Kato99] H. Kato e M. Billinghurst. Marker Tracking and HDM Calibration for a Video-Based Augmented Reality Conferencing System. *International Workshop on Augmented Reality (IWAR)*, San Francisco, USA, 1999.

[OCV07] Open CV Website, Junho de 2007.

<<http://www.opencv.org>>

[OGL07] Open GL Website, Junho de 2007.

<<http://www.opengl.org>>

[P3D07] Photo 3D Website, Junho de 2007.

<<http://www.photo3d.com>>

[Rolland01] J. P. Rolland, L. D. Davis e Y. Baillot. A Survey of Tracking Technology for Virtual Environments. *Fundamentals of Wearable Computers and Augmented Reality*, LEA Publishers, 2001

[Simon02] G. Simon e M-O. Berger. Reconstructing while Registering: a Novel Approach for Markerless Augmented Reality. *ISMAR '02 Conference Proceeding*, 2002.