



**FCTUC DEPARTAMENTO DE ENGENHARIA ELECTROTÉCNICA E DE COMPUTADORES**  
**FACULDADE DE CIÊNCIAS E TECNOLOGIA**  
**UNIVERSIDADE DE COIMBRA**

# **CALIBRAÇÃO DE CÂMARAS NÃO-CENTRAIS USANDO "THIN-PLATE-SPLINES"**

AUTORIA:

**Fernando Emanuel Saraiva Paulino**

PRESIDENTE DO JÚRI

**Professor Doutor Jorge Manuel Moreira de Campos Pereira Batista**

ORIENTADOR

**Professor Doutor Helder de Jesus Araujo**

VOGAL

**Professor Doutor Jorge Nuno de Almeida e Sousa Almada Lobo**

Coimbra  
Setembro de 2011



# AGRADECIMENTOS

Durante este longo percurso, várias foram as pessoas que contribuíram para levar este trabalho a bom termo. Recordo todos aqueles que me apoiaram ao longo destes anos, amigos, colegas e claro todos os Professores que contribuíram para a minha formação académica.

Em primeiro lugar, gostaria de agradecer ao meu orientador Professor Doutor Helder de Jesus Araujo, por todo o tempo que disponibilizou, pelos ensinamentos e pelos conhecimentos transmitidos.

Agradeço ao Pedro Miraldo, pela insistência e dedicação, pelas sugestões e pelas correcções com que contribuiu para a realização desta dissertação e acima de tudo por ser um grande amigo.

Agradeço também aos meus colegas e amigos por todo o apoio emocional prestado, em especial à Inês e à Ana que, ao longo deste ano, sempre tiveram as palavras que eu precisava ouvir.

Por fim, gostaria de agradecer à minha família, com especial atenção para meus pais e irmãos que, de diversas formas me apoiaram e sempre acreditaram em mim.

# ÍNDICE

<b>Lista de Figuras</b>	<b>iv</b>
<b>Lista de Tabelas</b>	<b>v</b>
<b>Lista de Acrónimos</b>	<b>vi</b>
<b>Resumo</b>	<b>vii</b>
<b>Abstract</b>	<b>viii</b>
<b>Capítulo 1</b>	<b>1</b>
1.1. Introdução	1
1.2. Notações	2
1.3. Organização	3
<b>Capítulo 2 - Estado de arte</b>	<b>4</b>
2.1. Modelos de Câmara Paramétricos	4
2.1.1. <i>Pinhole</i>	4
2.1.2. <i>Sistema Catadioptrico com Espelho Esférico</i>	6
2.1.3. <i>Sistema Catadioptrico Multi-Perspectivo</i>	8
2.2. Modelos de Câmara Genéricos	9
<b>Capítulo 3 - Representação paramétrica do modelo genérico de imagem</b>	<b>13</b>
3.1. Interpolação	13
3.1.1. <i>Radial Basis Functions</i>	14
3.2. Coordenadas de Plücker	15
3.3. Representação Paramétrica do Modelo Genérico de Imagem	16
3.4. Calibração	20

<b>Capítulo 4 - Um ponto no mundo, vários pontos na imagem</b>	<b>23</b>
4.1. O problema das multi correspondências em modelos de imagem	23
<b>Capítulo 5 - Implementação Prática</b>	<b>25</b>
5.1. Calibração	25
5.1.1. <i>Aquisição do dataset</i>	27
5.1.2. <i>Calibração</i>	27
5.1.3. <i>Resultados</i>	28
5.2. Algoritmo Computacional	29
<b>Capítulo 6 - Resultados</b>	<b>31</b>
<b>Capítulo 7 - Conclusões/Trabalho Futuro</b>	<b>40</b>
<b>Referências Bibliográficas</b>	<b>42</b>

# LISTA DE FIGURAS

<b>Figura 1</b> - Esquema do modelo de imagem perspectivo	5
<b>Figura 2</b> - Esquema do modelo de imagem catadioptrico com espelho esférico	7
<b>Figura 3</b> - Esquema do modelo de imagem catadioptrico com dois espelhos planos	8
<b>Figura 4</b> - Modelo genérico de imagem	9
<b>Figura 5</b> - Representação de uma superfície de incidência	10
<b>Figura 6</b> - Esquema representativo de um <i>Raxel</i>	12
<b>Figura 7</b> - Imagem exemplificativa do mapeamento de pixels para rectas 3D	17
<b>Figura 8</b> - Distância entre duas rectas incidentes no mesmo ponto 3D	24
<b>Figura 9</b> - Xadrez ( <i>A</i> ) e <i>plate</i> ( <i>B</i> ) utilizados	25
<b>Figura 10</b> - Sistema de medição (IR) <i>Optotraker</i>	26
<b>Figura 11</b> - Disposição do sistema de imagem catadioptrico com dois espelhos planos	28
<b>Figura 12</b> - Referencial utilizado	29
<b>Figura 13</b> - Fase de testes com diferentes distribuições de pontos de controlo	31
<b>Figura 14</b> - Fase de testes com as rectas obtidas na calibração	32
<b>Figura 15</b> - Modelo de imagem projectivo	34
<b>Figura 15a)</b> - Projectivo representação 2D	34
<b>Figura 15b)</b> - Projectivo representação 3D	34
<b>Figura 15c)</b> - Projectivo resultados	34
<b>Figura 16</b> - Modelo de imagem catadioptrico com espelho esférico	35
<b>Figura 16a)</b> - Esférico representação 2D	35
<b>Figura 16b)</b> - Esférico representação 3D	35
<b>Figura 16c)</b> - Esférico resultados	35
<b>Figura 17</b> - Modelo de imagem catadioptrico com dois espelhos planos	36
<b>Figura 17a)</b> - Multiperspectivo representação 2D	36
<b>Figura 17b)</b> - Multiperspectivo representação 3D	36
<b>Figura 17c)</b> - Multiperspectivo resultados	36
<b>Figura 18</b> - Estimação da intersecção de rectas	37
<b>Figura 18a)</b> - Intersecção de quatro pares de rectas	37
<b>Figura 18b)</b> - Estimação das intersecções	37
<b>Figura 18c)</b> - Estimação da intersecção de rectas com resultados	37

# LISTA DE TABELAS

<b>Tabela 1</b> - Tipos de <i>Radial Basis Functions</i>	14
<b>Tabela 2</b> - Média e desvio padrão valores obtidos nos testes práticos	38
<b>Tabela 3</b> - Média e desvio padrão na estimação das intersecções das rectas geradas	38
<b>Tabela 4</b> - Número de pontos utilizados	39

# LISTA DE ACRÓNIMOS

RBF	<i>Radial Basis Function</i>
SVD	<i>Single Value Decomposition</i>
MQ	Multiquadricas
TPS	<i>Thin-Plate-Splines</i>
CP	Centro de Projecção
LED	<i>Light-emitting Diode</i>
IR	<i>Infrared</i>



# RESUMO

A calibração de câmaras foi, desde sempre, um dos maiores desafios da visão por computador. Os métodos analíticos tornaram-se um dos limites a ultrapassar, pela sua falta de adaptação à complexidade, cada vez maior, dos modelos de câmaras não centrais.

No trabalho apresentado estudamos o estado da arte de alguns modelos de câmara, paramétricos e não paramétricos, suas características, vantagens e desvantagens. Procedemos à descrição do método proposto por Miraldo *et al.* [1] e apresentamos o trabalho prático desenvolvido que lhe serviu de suporte. Além das *Radial Basis Functions* usadas alargamos o conjunto de testes para incluir *thin-plate-splines*.

Relacionado com o conceito de vários *viewpoints*, abordamos a questão “um ponto no mundo, dois ou mais pontos na imagem”. A sua implementação prática é realizada e é feito um estudo sobre os fundamentos teóricos envolvidos. Esta abordagem permite uma estimação mais precisa do ponto incidente, quando um ponto no mundo é visto por mais do que um ponto na imagem.

**Palavras-chave:** Câmaras não-centrais, *thin-plate-splines*, modelo de imagem genérico, *radial basis functions*, calibração, paramétrico.

# ABSTRACT

Camera calibration has been one of the greatest challenges for the computer vision community. Most camera models are limited by their analytic nature, which is very restrictive. It isn't possible to make adjustments to these models, that allow them to surpass this limitation and be able to adapt to the non central camera models.

In this dissertation we study the state of art of some camera models, both parametric and non-parametric and we emphasize their advantages and disadvantages. We give special importance to Miraldo *et al.* [1] method, focusing on their theoretic basis. We also present our real data tests made to support the method's ground true. In these tests we include the Radial Basis Functions used before. We also tested for thin-plate-splines, which don't have shape parameter.

We approach the special situation "one 3D point, two or more image points", with both theoretical and practical resolutions. This method allows an accurate estimation of the incident point in systems with multiple viewpoints.

**Keywords:** Non central camera, radial basis functions, thin-plate-splines, general imaging model, calibration, parametric.

# CAPÍTULO 1

## 1.1. Introdução

Uma câmara é um sistema que permite o mapeamento de pontos do mundo em pontos na imagem, ou seja, uma projecção de pontos 3D para 2D que é feita através de uma recta. Dependendo da complexidade da câmara, o mapeamento varia e pode tornar-se de tal forma complexo que pode não existir lógica matemática para o compreender e replicar. Chama-se calibração a este mapeamento, ao conhecimento dos parâmetros da câmara e à forma como esta se comporta quando entra em contacto com a luz para a formação da imagem. Grande parte das câmaras existentes podem ser representadas por modelos paramétricos, de simples aplicação e com resultados bastante precisos. No entanto, muitas vezes os modelos são feitos à medida, sem que exista a possibilidade de uma adaptação para outros tipos de câmara. São modelos fechados e geometricamente rígidos onde não há lugar para simples mudanças. Existem também modelos genéricos aplicáveis a todos os modelos de câmara [1, 2, 3, 4, 5], mais abrangentes e com vantagens de adaptação, onde a câmara pode ser tratada como um sistema do qual não sabemos qualquer detalhe, uma *black-box*. No entanto, ao contrário dos modelos paramétricos, nestes existe o requisito de mapear todos os *pixels* individualmente. Em [2, 3, 4, 5] os autores defendem modelos genéricos não paramétricos que, apesar de poderem calibrar diversos modelos de câmara, por serem discretos, são computacionalmente muito pesados.

Como veremos mais à frente, a conjugação de paramétrico e genérico pode ser considerada uma enorme vantagem no processo de calibração. Ao fundir as qualidades de ambos, é possível obter um método relativamente simples, preciso e abrangente que permite calibrar vários tipos de câmaras. O trabalho experimental apresentado serviu de suporte à fundamentação teórica de Miraldo *et al.* [1]. Pretendemos acrescentar aos resultados obtidos um tipo de *RBF* que não possui parâmetro de forma. A calibração é um dos maiores desafios para a investigação, onde o número de imagens e a quantidade de pontos pesam na computação, precisão e conseqüente rapidez de processos, possibilitando a recuperação das propriedades métricas geralmente chamadas de parâmetros internos. Muitos modelos até hoje de difícil solução, com cálculos dantescos, podem ser facilmente representados pelo modelo de imagem genérico.

## 1.2. Notações

As matrizes são representadas por letras maiúsculas a negrito (exemplo:  $\mathbf{A} \in \mathbb{R}^{n \times m}$ ,  $n$  linhas e  $m$  colunas). Os vectores são representados por letras minúsculas a negrito (exemplo:  $\mathbf{a} \in \mathbb{R}^n$ ,  $n$  elementos). Se nada for dito, um vector é considerado uma coluna. Letras minúsculas são consideradas elementos de uma dimensão (exemplo:  $a$ ). Por defeito, a coluna  $j$  da matriz  $\mathbf{A}$  é escrita como  $\mathbf{a}_j$ . O elemento  $j$  do vector  $\mathbf{a}$  é escrito como  $a_j$ . O elemento  $\mathbf{A}$  da linha  $i$  e da coluna  $j$  é representado por  $a_{i,j}$ . Letras maiúsculas indicam constantes de uma dimensão (exemplo:  $A$ ).

O espaço projectivo é representado por  $P^n$  (espaço Euclideano de dimensão  $n$ ). Um ponto  $\mathbf{x}$  em  $P^n$  pode ser escrito em coordenadas homogéneas em  $\mathbb{R}^{n+1}$  como  $\mathbf{x} = (x_0, x_1, x_2, \dots, x_n)$  e pode ser convertido para coordenadas não homogéneas por  $\mathbf{x} = (x_1/x_0, x_2/x_0, \dots, x_n/x_0)$ .

É usado  $\mathbb{R}$  após uma matriz ou um vector para indicar que o mesmo está representado a menos de um factor de escala.

Seja  $\mathbf{U} \in \mathbb{R}^{m \times n}$  e  $\mathbf{V} \in \mathbb{R}^{k \times l}$  e a equação

$$\mathbf{UXV}^T = \mathbf{C} \quad (1)$$

onde  $\mathbf{X} \in \mathbb{R}^{n \times l}$  é a incógnita. É possível escrever a equação anterior da seguinte forma

$$(\mathbf{V} \otimes \mathbf{U})\text{vec}(\mathbf{X}) = \text{vec}(\mathbf{C}) \quad (2)$$

onde  $\otimes$  é o produto de Kronecker de  $\mathbf{U}$  e  $\mathbf{V}$ , com  $(\mathbf{V} \otimes \mathbf{U}) \in \mathbb{R}^{mk \times nl}$ , e  $\text{vec}(\mathbf{X})$  é um vector de dimensão  $nl$  formado pela vectorização em colunas de  $\mathbf{X}$ .

Para uma matriz  $\mathbf{A}$ , o  $\mathbf{x}$  que minimiza  $\mathbf{Ax}$  é dado pela última coluna de  $\mathbf{V}$ , onde [6],

$$\mathbf{A} = \mathbf{UDV}^T \quad (3)$$

e,  $[\mathbf{U}, \mathbf{D}, \mathbf{V}] = \text{svd}(\mathbf{A})$  (single value decomposition).

Uma solução possível do sistema é  $\mathbf{x}=\mathbf{0}$ , no entanto esta não tem interesse dado que procuramos uma solução diferente de zero<sup>1</sup>. Seja  $\mathbf{x}$  um vector solução e  $k$  um escalar, então todo o  $k\mathbf{x}$  satisfaz a equação  $\mathbf{Ax}=\mathbf{0}$ . Aplicando a restrição  $\|\mathbf{x}\|=1$  no sentido de minimizar  $\|\mathbf{Ax}\|$ , ou seja (equação (3)),  $\|\mathbf{UDV}^T\mathbf{x}\|$ , sabemos que  $\|\mathbf{UDV}^T\mathbf{x}\|=\|\mathbf{DV}^T\mathbf{x}\|$  pelas propriedades das transformações ortonormais [6] e  $\|\mathbf{x}\|=\|\mathbf{V}^T\mathbf{x}\|$ . Então, podemos minimizar  $\|\mathbf{DV}^T\mathbf{x}\|$  com a condição  $\|\mathbf{V}^T\mathbf{x}\|=1$  que passa por escrever  $\mathbf{y}=\mathbf{V}^T\mathbf{x}$ . Com  $\|\mathbf{Dy}\|$  sujeito à restrição  $\|\mathbf{y}\|=1$ . Sabendo que  $\mathbf{D}$  é a matriz diagonal dos valores singulares por ordem descendente obtemos a solução  $\|\mathbf{y}\|=[0,0,\dots,0,1]^T$ , uma matriz de zeros com um 1 no último elemento. Sabendo que  $\text{inv}(\mathbf{V}^T)=\mathbf{V}$  concluímos que a solução é  $\mathbf{x}=\mathbf{Vy}$ , ou seja, a última coluna de  $\mathbf{V}$ .

### 1.3. Organização

A dissertação está dividida em sete capítulos. Inicialmente, no primeiro capítulo, é feita uma breve introdução onde tratamos o problema, a motivação que nos levou a desenvolver este trabalho e algumas notações para facilitar a descrição dos métodos. Em seguida, no segundo capítulo, é apresentado o estado de arte: alguns modelos de câmara paramétricos e não paramétricos. São abordados os três sistemas de câmaras usados nos resultados e dois modelos genéricos não paramétricos. No terceiro capítulo é explicado o método no qual se baseia este trabalho. Começamos por explicar a base matemática para depois se aprofundar o conceito de calibração. No quarto capítulo é abordado o conceito de múltiplos pontos na imagem poderem ver o mesmo ponto no mundo e suas repercussões na representação do modelo genérico de imagem [1]. No capítulo quinto é demonstrada a componente prática, onde se prova a viabilidade do método e se descreve passo a passo os algoritmos a usar na sua implementação. Em seguida, no capítulo sexto, são comentados os resultados obtidos. Por fim, é apresentada uma conclusão do trabalho efectuado, correspondente ao capítulo sétimo.

---

<sup>1</sup> Para uma matriz  $\mathbf{A}$  de dimensão  $m \times n$  com característica menor que  $n$  existem infinitas soluções diferentes de zero.

# CAPÍTULO 2 - ESTADO DE ARTE

## **2.1. Modelos de Câmara Paramétricos**

Uma câmara permite o mapeamento do mundo 3D para uma imagem 2D. Os vários modelos de câmara englobam diversas taxonomias e anatomias com respectivas vantagens e desvantagens. Geralmente, o mapeamento de uma câmara pode ser representado por uma matriz de projecção 3x4 denominada P, que transforma coordenadas 3D em coordenadas no plano da imagem. A matriz referida permite conhecer parâmetros intrínsecos da câmara tais como, a distância focal, o ponto principal, factores de escala ou extrínsecos como o seu centro.

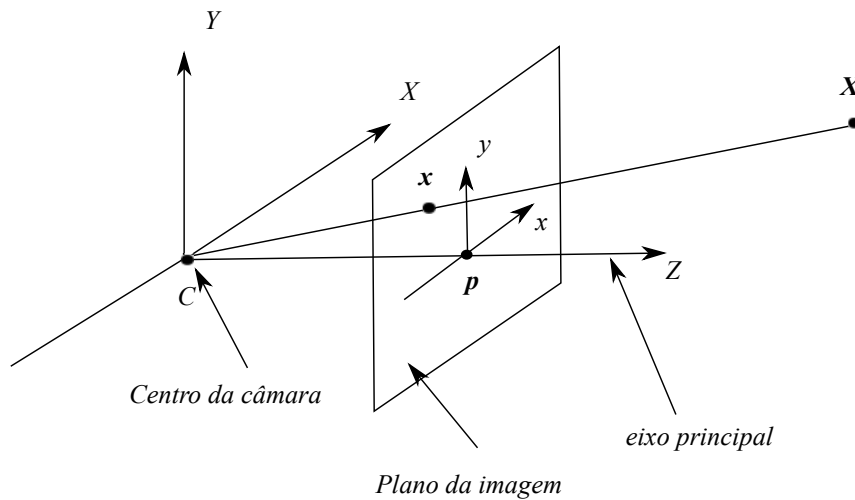
Existem vários tipos de câmara diferentes entre si, no entanto a base para o cálculo das suas características geométricas está incorporada na representação matricial, em que os valores são geralmente de fácil aquisição.

O caso das câmaras centrais pode ser considerado o mais estudado e de abordagem mais simples. No entanto, devido a diversos factores como amplitude do campo de visão ou mesmo do material de que é feita a lente, as câmaras não centrais tornaram-se mais comuns. Ao contrário das câmaras centrais, as não centrais, na generalidade, não são passíveis de estimação paramétrica para o cálculo das suas características intrínsecas e extrínsecas.

Conjugações de espelhos e câmaras permitem criar um sistema, onde existe um centro virtual que não se encontra em nenhuma das diversas câmaras. Globalmente, este centro reúne as características de um sistema central, ou seja, os raios de projecção passam todos num e apenas num só ponto. Os sistemas podem ser constituídos por espelhos com forma esférica, elíptica, hiperbólica ou parabólica. Se a câmara for colocada nos *foci* dos espelhos (à excepção do parabólico) o sistema pode ser considerado central [7]. No entanto, as restrições para que a combinação de espelhos e câmaras resulte em sistemas aproximadamente centrais é elevada, por isso o resultado deste tipo de sistemas é quase sempre um sistema não central.

### **2.1.1. Pinhole**

Neste modelo é considerada a projecção de pontos do espaço num plano. Assumindo o centro da projecção como a origem do referencial de coordenadas e colocando o eixo Z normal ao plano temos a seguinte representação.



**Figura 1** - Esquema representativo do modelo de câmara perspectivo.

Então, um ponto de coordenadas  $\mathbf{X} = \begin{pmatrix} X & Y & Z \end{pmatrix}^T$  será mapeado no ponto onde a recta que une  $C$  e  $X$  intersecta o plano da imagem. A distância da origem do referencial até ao ponto  $p$ , ponto principal, é denominada de distância focal  $f$ .

O factor essencial nestas câmaras é que todos os raios se intersectam num só ponto [6], no entanto, este modelo, apesar de ser simples de analisar é demasiado restritivo.

Assumindo um ponto na imagem  $\mathbf{x} = (x, y)$  correspondente ao ponto  $\mathbf{X}$  no mundo podemos dizer que,

$$\mathbf{x} = f \frac{\mathbf{X}}{\mathbf{Z}} \qquad y = f \frac{Y}{Z} \qquad (4)$$

Sabemos também que é possível efectuar uma transformação de corpo rígido do referencial da câmara para o referencial do mundo caso estes não sejam coincidentes.

Então,

$$\mathbf{P}_C = \begin{pmatrix} X_C \\ Y_C \\ Z_C \end{pmatrix} = \mathbf{R}\mathbf{P}_W + \mathbf{T} = \begin{pmatrix} r_1 & r_2 & r_3 \\ r_4 & r_5 & r_6 \\ r_7 & r_8 & r_9 \end{pmatrix} \mathbf{P}_W + \begin{pmatrix} t_x \\ t_y \\ t_z \end{pmatrix} \quad (5)$$

Considerando o  $\mathbf{p}$  como sendo o centro da imagem (Figura 1), podemos calcular as coordenadas do ponto  $\mathbf{x}$  em função desse centro,

$$x_f = k_x x + C_x \quad y_f = k_y y + C_y \quad (6)$$

Combinando as equações (5) e (6) obtemos o seguinte sistema na forma matricial.

$$\begin{pmatrix} x_f \\ y_f \\ 1 \end{pmatrix} = \begin{pmatrix} P_1 & P_2 & P_3 & P_4 \\ P_5 & P_6 & P_7 & P_8 \\ P_9 & P_{10} & P_{11} & P_{12} \end{pmatrix} \begin{pmatrix} X_W \\ Y_W \\ Z_W \\ 1 \end{pmatrix} \quad (7)$$

C representa o referencial da câmara e W representa o referencial do mundo.

A matriz  $\mathbf{P}$  permite então transformar pontos do mundo em pontos da imagem.

As câmaras onde os raios não se intersectam num só ponto são denominadas câmaras não centrais.

### **2.1.2. Sistema Catadioptrico com Espelho Esférico**

Segundo Agrawal et al. [7] é possível fazer uma adaptação para alguns casos de câmaras não centrais permitindo o uso de equações paramétricas. É o caso do sistema catadioptrico com espelho esférico. Os autores demonstram que se a câmara for colocada no eixo de simetria do espelho, é possível obter uma solução fechada que permite o mapeamento de pontos na imagem para rectas no mundo.

Os sistemas catadioptricos permitem obter grandes campos de visão, no entanto, a qualidade de imagem fica degradada devido às características da câmara, do espelho, da combinação de câmaras ou da combinação de espelhos. Existem diversos casos em que é possível tornar estes sistemas calibráveis, onde com configurações bastante restritas de câmaras é possível obter uma aproximação



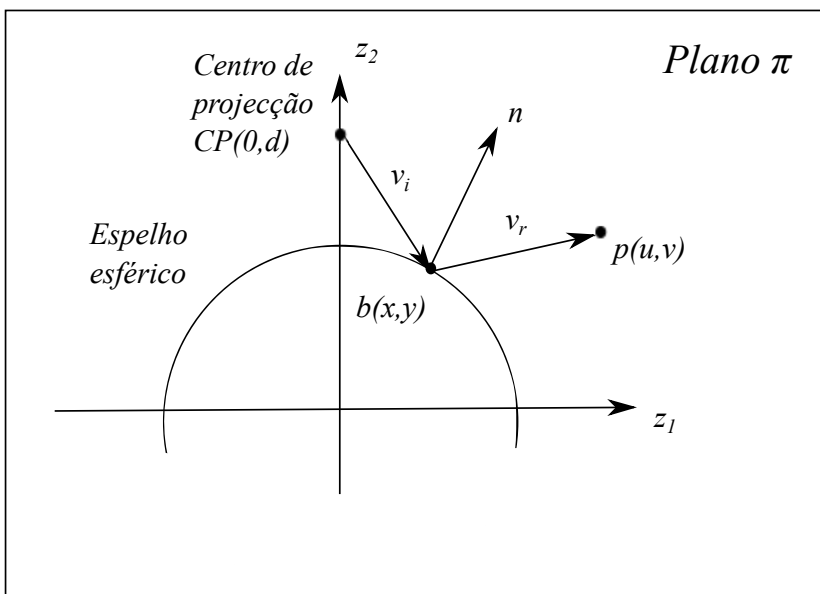
a um sistema central. A condição essencial é a câmara encontrar-se no eixo de simetria do espelho para que o número de incógnitas seja o mais baixo possível.

Segundo o mesmo autor, para um dado  $\mathbf{C}_p = \begin{bmatrix} 0 & 0 & -d \end{bmatrix}^T$  e um espelho esférico de raio  $r$  e para um ponto  $\mathbf{q}$  de uma imagem, seja  $\mathbf{s} = \mathbf{K}^{-1}\mathbf{q}$  a direcção da recta, onde  $\mathbf{K}_{3 \times 3}$  é a matriz de calibração interna da câmara. Então, conforme é possível observar na Figura 2, os pontos de intersecção  $\mathbf{b}$  com o espelho são dados por

$$\mathbf{b} = \mathbf{C}_p + \mathbf{s} \frac{ds_3 \pm \sqrt{d^2 s_3^2 - (d^2 - r^2)(\mathbf{s}^T \mathbf{s})}}{\mathbf{s}^T \mathbf{s}} \quad (8)$$

É importante ter em conta que  $\mathbf{b}^T \mathbf{b} = r^2$ , pois o centro de projecção é coincidente com o centro da esfera e a normal de  $\mathbf{b}$  é  $\mathbf{b}/r$ . Sabendo que  $\mathbf{v}_i = \mathbf{b} - \mathbf{C}_p$ , em que  $\mathbf{v}_i$  é o vector incidente, podemos calcular o vector reflectido  $\mathbf{v}_r$  que é dado por

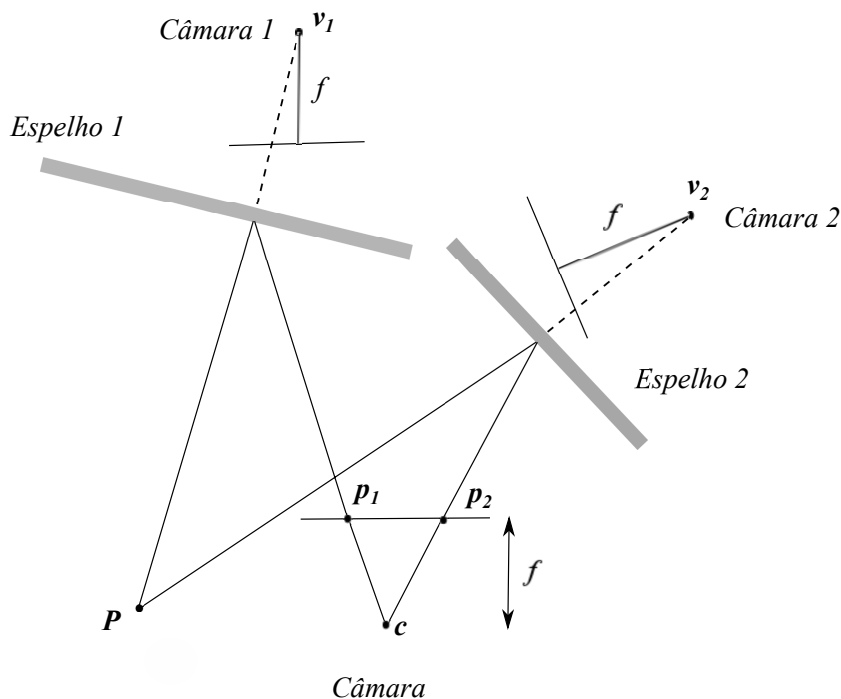
$$\mathbf{v}_r = (\mathbf{b} - \mathbf{C}_p) - 2\mathbf{b}(\mathbf{b}^T (\mathbf{b} - \mathbf{C}_p)) / r^2 = -\mathbf{b} - \mathbf{C}_p + 2\mathbf{b}(\mathbf{b}^T \mathbf{C}_p) / r^2 \quad (9)$$



**Figura 2** - Esquema representativo do modelo catadióptrico com espelho esférico e câmara centrada no eixo principal [7]. O plano  $\pi$  é definido pelo ponto no espelho  $b(x,y)$ , o centro de projecção  $CP(0,d)$  e a origem do referencial.

### 2.1.3. Sistema Catadioptrico Multi-Perspectivo

O sistema catadioptrico multi-perspectivo traduz-se na utilização de dois ou mais espelhos planos conjugados com uma câmara projectiva (Figura 3). Cada espelho terá um centro de projecção associado, que não é mais do que uma translação e rotação virtual da câmara, para um ponto atrás do espelho e que pertence à normal do seu plano. O sistema corresponde ao modelo projectivo quando a análise é feita individualmente.



**Figura 3** - Esquema exemplificativo de um modelo catadioptrico com dois espelhos planos.  $\mathbf{V}_i$  são centros de câmaras virtuais,  $\mathbf{P}_i$  são pontos principais e  $f$  é a distância focal.  $\mathbf{P}$  é o ponto no mundo [8].

Partindo da posição da câmara, podemos obter as posições  $\mathbf{V}_i$  através de uma transformação de corpo rígido  $\mathbf{D}$ .

$$\mathbf{v}_i = \mathbf{D}_i \mathbf{c} \quad (10)$$

em que  $\mathbf{D}$  é dado por,

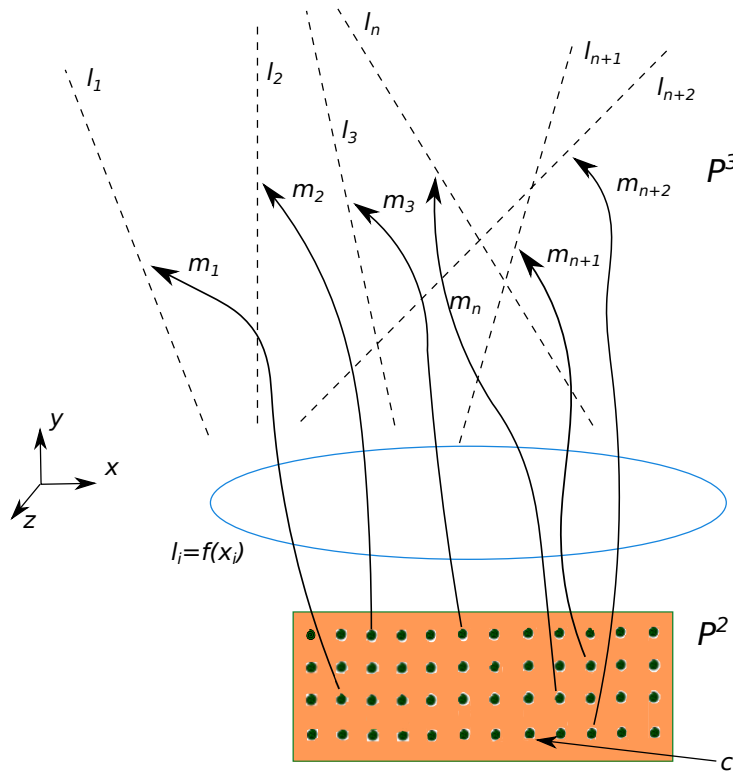
$$\mathbf{D}_i = \begin{pmatrix} \mathbf{I} - 2\mathbf{n}_i\mathbf{n}_i^T & 2d_i\mathbf{n}_i \\ 0 & 1 \end{pmatrix} \quad (11)$$

onde  $\mathbf{n}$  é a normal ao plano espelho e  $d$  é a distância à câmara  $\mathbf{c}$ .

## 2.2. Modelos de Câmara Genéricos

Há muito que o modelo projectivo é usado na visão por computador. No entanto, as limitações associadas à linearidade conjugadas com a multiplicidade de modelos não lineares, levou a que surgissem defensores da utilização de um modelo genérico de imagem.

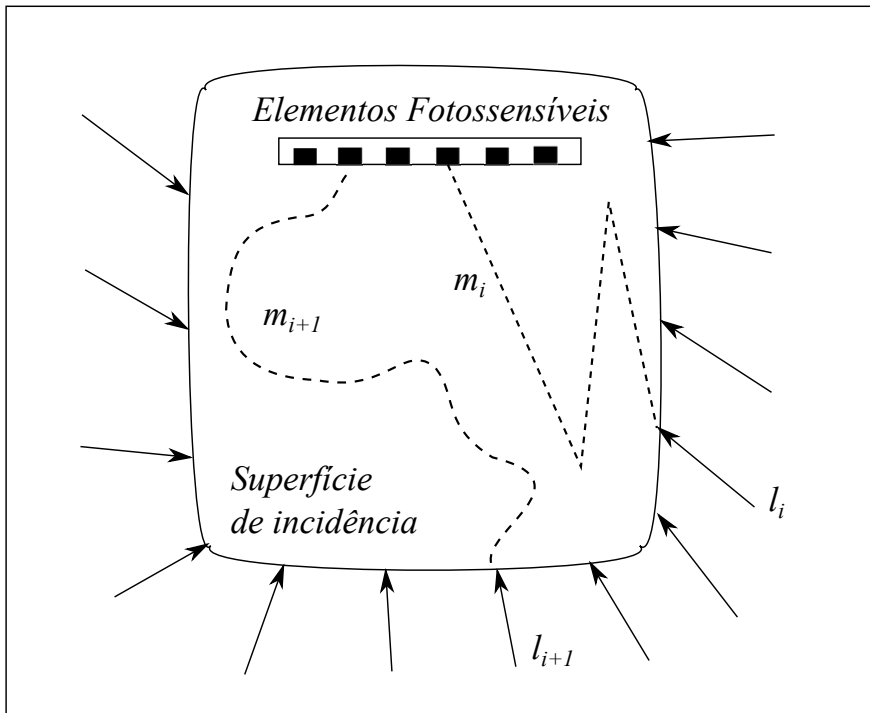
Os sistemas generalizados de câmaras permitem adaptar e direccionar os sistemas, ao encontro das características que o utilizador pretende. A conjugação de modelos de câmaras centrais e não centrais possibilita um maior campo de visão e um aumento significativo da resolução, no entanto, a complexidade da sua calibração também aumenta. Encontrar lineariedade ou uma aproximação computacionalmente exequível, permite solucionar a calibração destes sistemas baseados muitas vezes na heterogenia de câmaras.



**Figura 4** - modelo genérico de imagem.  $m_i$  representa o mapeamento de *pixels* de  $P^2$  em rectas em  $L^3$  (que representam os parâmetros geométricos dos *raxels*). Consideramos que o mapeamento é relativamente suave e por isso pode ser representado por uma função de vectores  $\mathbf{f}(\mathbf{x}): P^2 \rightarrow L^3$ .  $c_i$  são denominados pontos de controlo.

O modelo genérico de imagem [2, 3] levou à junção dos diversos modelos de câmara, que eram em termos de características, muitas vezes extremamente distantes. Este conceito ignora por completo o

percurso que a luz faz quando passa por lentes e espelhos do sistema de imagem, em vez disso, centraliza-se na relação entre *pixels* e a região do espaço que afecta o sensor da câmara (Figura 5).



**Figura 5** - Representação esquemática da superfície de incidência e do dispositivo fotossensível. Assumimos que existe uma correspondência entre o *pixel*  $i$  e o raio  $\mathbf{l}_i$ .  $m_i$  representa o mapeamento de *pixels* de  $P^2$  em rectas em  $L^3$ .

O mapeamento dos raios incidentes pode ser visto como uma superfície facilmente representada pelos receptores do elemento fotossensível, traduzindo-se numa relação em que o raio  $\mathbf{l}_i$  corresponde ao *pixel*  $i$ , descurando por completo o percurso que o raio faz até chegar ao *pixel*. Importa salientar que os raios possuem diversas propriedades (geométricas, radiométricas e ópticas [2, 3]), mas que em parte podem ser ignoradas [4, 5]. Desta forma, um sistema de imagem que seja desconhecido pode ser tratado de forma simples, pois o importante são os *pixels* da imagem e as rectas no espaço correspondentes, não importa a forma como a imagem é adquirida, o campo de visão ou se a câmara tem características não lineares.

Diversos autores (Nayar e Grossberg, Sturm e Ramalingam) têm vindo a convergir nesse sentido para simplificar os processos de calibração e análise de imagens, fornecidas pelos vários tipos de câmaras. Uma das abordagens que mais modelos de câmara enquadra, o modelo de imagem generalizado [2, 3] é baseado em *raxels* (Figura 6) e aproxima-se da realidade computacionalmente exequível. Os *raxels* incluem propriedades geométricas, radiométricas e ópticas [2, 3], no entanto, as propriedades radiométricas e ópticas podem ser eliminadas para que o modelo fique mais simples [4, 5]. Um *raxel* não é mais do que um conjunto de parâmetros geométricos, radiométricos e ópticos

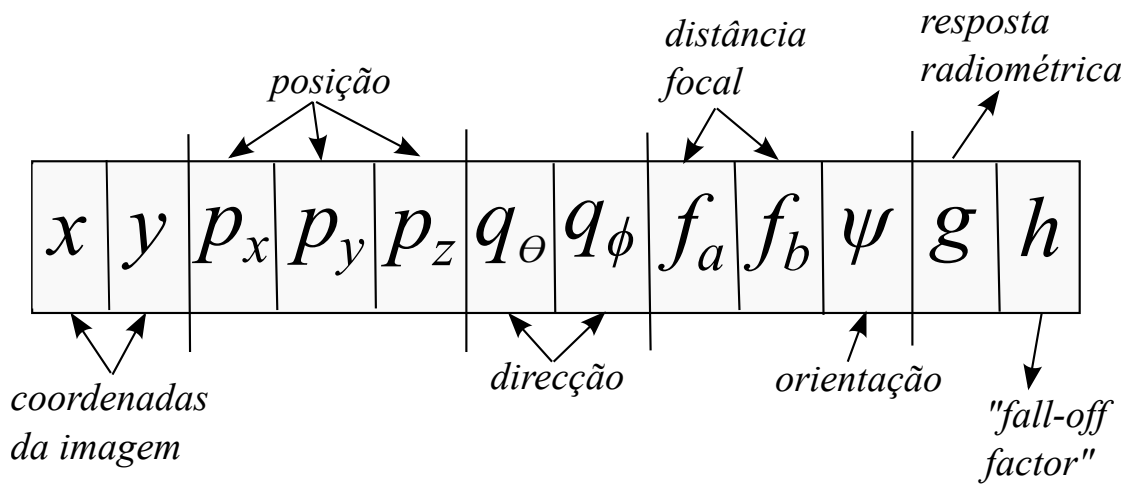
que ficam associados a um *pixel*. Este inclui, na versão de Nayar e Grossberg coordenadas do *pixel* na imagem, um vector associado com a sua posição e direcção, assim como parâmetros radiométricos e ópticos. Através da representação com *raxels* é relativamente simples englobar sistemas de várias câmaras, centrais ou não centrais, assim como sistemas catadioptricos com e sem distorção. Na abordagem de Grossberg e Nayar [2, 3] é proposto um método que permite a estimação dos parâmetros para o modelo de imagem generalizado. Partindo do pressuposto que um raio  $\mathbf{l}_i$  corresponde ao *pixel*  $i$ , que este raio intersecta dois planos, separados por uma distância conhecida, em pontos dos planos conhecidos, é possível saber a direcção do raio para cada *pixel* e assim calcular os parâmetros do *raxel*. Após este processo, ao efectuar uma rotação da câmara é possível fazer uma calibração baseada nos parâmetros dos *raxels*. O método necessita de uma recolha exaustiva de imagens.

O método proposto por Ramalingam *et al.* [4] e Sturm e Ramalingam [5] difere do anterior devido à exclusão da informação radiométrica, limitando o mapeamento à relação entre raios e *pixels*. Enquanto que no método de Nayar e Grossberg eram necessárias duas ou mais imagens de um objecto conhecido, sendo também necessário saber a posição da câmara em cada aquisição, no caso de Sturm e Ramalingam bastam três ou mais imagens dos objectos de calibração, sem que seja necessário saber o posicionamento da câmara.

Mais tarde, os mesmos autores [9] propuseram ainda um modelo minimalista mais robusto e resistente ao ruído quando comparado com os métodos anteriores.

Qualquer um dos três exemplos referidos é discreto, não paramétrico e usa *raxels* ou variantes de *raxels* para efectuar a calibração.

Neste trabalho, pretendemos fazer uma extensão aos testes feitos ao método proposto por Miraldo *et al.* [1], onde a complexidade do modelo de imagem generalizado é reduzida pelo pressuposto que a relação entre *pixel* e recta varia suavemente e de forma controlada, permitido definir uma representação paramétrica aproximada (Figura 4). Além dos testes já efectuados [1], pretendemos aplicar *Radial Basis Functions* sem parâmetro de forma, como é o caso das *Thin Plate Splines*. Com o método do referido autor, podem ser obtidos resultados que não dependem da resolução da imagem e o número de incógnitas é largamente reduzido permitindo a simplificação dos cálculos a



**Figura 6** - Cada *raxel* possui os parâmetros indicados na imagem. Existem variantes do esquema proposto por Nayar Grossberg que eliminam as componentes radiométrica e óptica.

efectuar. Nos métodos propostos por Nayar e Grossberg, Ramalingam e Sturm [2, 3, 4, 5] são sempre necessários pelo menos dois pontos no mundo para o mesmo *pixel*, enquanto que na nova abordagem [1] apenas é necessário um ponto no mundo para um ponto na imagem.

# CAPÍTULO 3 - REPRESENTAÇÃO PARAMÉTRICA DO MODELO GENÉRICO DE IMAGEM

## 3.1. Interpolação

Quando se pretende calcular uma aproximação de um conjunto de pontos aleatórios é possível, utilizando uma função de interpolação, estimar valores não conhecidos com relativa precisão. A interpolação depende essencialmente da quantidade de pontos para análise, da sua dispersão e também da função de estimação usada.

Para estimar uma função  $f: \mathbb{R}^D \rightarrow \mathbb{R}$  através de um conjunto aleatório de pontos  $X = \{\mathbf{x}_i\} \subset \mathbb{R}^D$ , (sendo  $D$  um número natural) e um conjunto de pontos  $Y = \{y_i\}$ , onde o conjunto dos  $\mathbf{x}_i$  e  $y_i$  formam  $y = f(\mathbf{x}_i), y \in \mathbb{R}$  e  $x \in \mathbb{R}^n$ , é necessário calcular a função  $s: \mathbb{R}^D \rightarrow \mathbb{R}$  que satisfaz a equação

$$s(\mathbf{x}_i) = f(\mathbf{x}_i), \forall_i \quad (12)$$

Para resolver o problema podemos usar *Radial Basis Functions (RBF)* [1,10,11].

Para um determinado P, conjunto de pontos de treino  $\{\mathbf{x}_1, \dots, \mathbf{x}_P\}$ , a interpolação é dada por

$$s(x) = a_0 + \mathbf{a}_x^T \mathbf{x} + \sum_{i=1}^P \omega_i \phi(\|\mathbf{x} - \mathbf{x}_i\|) \quad (13)$$

onde  $a_0, \mathbf{a}_x$  e  $\omega_i$  são as incógnitas da interpolação que precisam de ser estimadas,  $\phi$  é a RBF e  $\|\mathbf{x} - \mathbf{x}_i\|$  é a norma de um vector. A função  $s(x)$  tem  $P+D+1$  graus de liberdade, no entanto apenas nos são fornecidas  $P$  equações através do conjunto de pontos  $\mathbf{X}$ . Sabemos também que para uma

função  $\phi$  positiva definida [10] as equações seguintes são verificadas [1,10,12] e o seu uso permite que a estimação das restantes incógnitas seja feita.

$$\sum_{i=1}^P \omega_i = 0 \quad \& \quad \sum_{i=1}^P \omega_i x_1^{(i)} = \dots = \sum_{i=1}^P \omega_i x_D^{(i)} = 0 \quad (14)$$

onde  $x_i^{(j)}$  é o elemento  $i$  da observação  $j$ .

### 3.1.1. Radial Basis Functions

Existem diversos tipos de base para as *Radial Basis Functions (RBF)* que podem ser divididos em dois tipos, com e sem parâmetros de forma. São exemplos do primeiro caso a *gaussiana* e a *multi quadrica*. É exemplo do segundo caso, ou seja, sem parâmetros de forma, as *thin-plate-splines*. No trabalho realizado foram usados os três exemplos referidos, no entanto, existem outros exemplos, apresentados na Tabela 1.

Tipos de RBF	Exemplo	Parâmetro de forma
Gaussiana	$\phi(r) = \exp(-\beta^2 r^2)$	Sim
Multi Quadricas	$\phi(r) = (\beta^2 + r^2)^{\frac{1}{2}}$	Sim
Multi Quadricas Inversa	$\phi(r) = \frac{1}{\sqrt{1 + (\beta r)^2}}$	Sim
Quadrica Inversa	$\phi(r) = \frac{1}{1 + (\beta r)^2}$	Sim
Poli harmónica (I)	$\phi(r) = r^k, k = 1, 3, 5, \dots$	Não
Poli harmónica (II)	$\phi(r) = r^k \ln(r), k = 2, 4, 6, \dots$	Não
<i>Thin-Plate-Spline</i>	$\phi(r) = r^2 \ln(r)$	Não

**Tabela 1** - Tipos de *radial basis functions* e alguns exemplos



De notar que as *thin-plate-splines* são um caso especial das *poli-harmónicas(II)* quando  $k=2$ .

### 3.2. Coordenadas de Plücker

As coordenadas de *Plücker* têm como base as coordenadas de *Grassmann*  $\Lambda^k \mathbb{R}^n$  com  $k=2$  e  $n=4$ , em que  $k$  são subespaços e  $n$  a dimensão de um vector [13]. As coordenadas de *Plücker* resultam do produto externo entre dois vectores de quatro dimensões, sendo o resultado obtido um vector de dimensão seis, que representa rectas em  $P^3$ . O vector resultante é formado por dois vectores de dimensão três,  $\mathbf{d}$  e  $\mathbf{m}$ , o vector direcção e o vector momento da recta, que passamos a designar por  $\mathbb{R}$ .

A recta  $\mathbb{R}$  pode ser definida como:

$$\mathbb{R} = \mathbf{x} \wedge \mathbf{w} = (l_{01}, l_{02}, l_{03}, l_{23}, l_{31}, l_{12}) \in \Lambda^2 \mathbb{R}^4 \subset \mathbb{R}^6 \quad (15)$$

onde a direcção é dada por  $\mathbf{d} = (l_{01}, l_{02}, l_{03})$  e o momento por  $\mathbf{m} = (l_{23}, l_{31}, l_{12})$ , com  $l_{ij} = x_i w_j - x_j w_i$ , em que  $e_{ij} = e_i \wedge e_j, e_i \in \mathbb{R}^4$ ,  $\mathbf{x}$  e  $\mathbf{w}$  são dois pontos em  $P^3$ , pertencentes à recta.

Pode ser provado que a equação anterior (15) é resultado do produto externo de dois vectores de dimensão quatro, e que por isso representam uma recta no espaço 3D, se e só se  $\mathbf{l}$  pertencer à quadrica de Klein [13].

$$\Omega(\mathbf{l}, \mathbf{l}) = l_{01}l_{23} + l_{02}l_{31} + l_{03}l_{12} = \langle \mathbf{d}, \mathbf{m} \rangle = 0 \quad (16)$$

A conjugação de todos os factores permite obter uma forma relativamente simples de representar rectas no espaço, que podem ser usadas para efectuar cálculos lineares e relativamente leves a nível computacional.

Através da direcção e do momento é possível construir um sistema que relaciona a incidência num ponto  $\mathbf{p}$  pertencente a  $P^3$  de uma recta  $\mathbb{R}$  pertencente a  $\Lambda^2 \mathbb{R}^4$ . A possibilidade de calcular computacionalmente a relação de incidência, permite obter,

$$\mathbf{Q}(\mathbf{p})\mathbf{l}\mathbb{R} = 0 \quad (17)$$

em que,

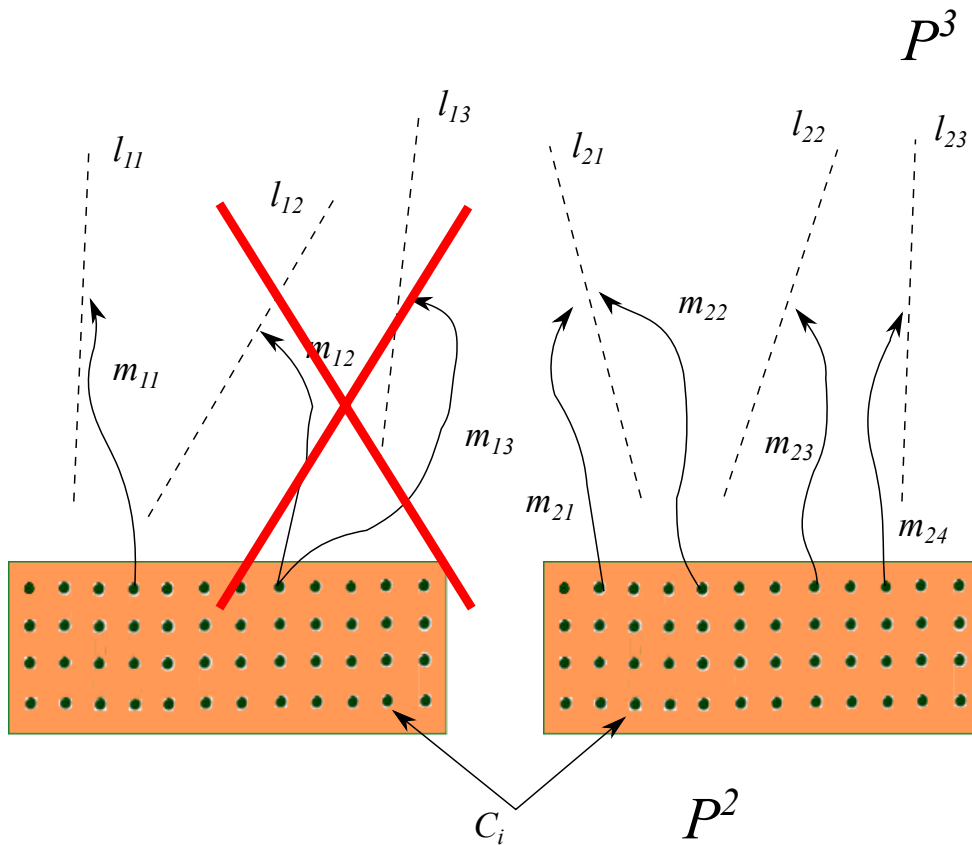
$$\mathbf{Q}(\mathbf{p}) = \begin{bmatrix} [\mathbf{p}]_x & -\mathbf{I} \\ \mathbf{0} & \mathbf{p}^T \end{bmatrix} \quad (18)$$

onde  $\mathbf{p}$  pertence a  $\mathbb{R}^3$ ,  $[\mathbf{a}]_x$  é uma matriz que lineariza o produto externo como  $[\mathbf{a}]_x \mathbf{b} = \mathbf{a} \times \mathbf{b}$  e  $\mathbf{I}$  é a matriz identidade 3 x 3.

### 3.3. Representação Paramétrica do Modelo Genérico de Imagem

Ambos os modelos paramétricos abordados [2, 3, 4, 5] são considerados discretos, ou seja, definem um mapeamento entre cada *pixel* da imagem e uma recta 3D no espaço. Como foi dito anteriormente (Secção 2.2), este mapeamento é baseado num conjunto de valores denominado *raxel*. Assim como em [1, 4, 5], considerando apenas a parte geométrica do *raxel*, ficamos com sete incógnitas que necessariamente precisam de ser estimadas. Logo, para uma imagem com o tamanho de N por M *pixels*, existem 7NM parâmetros para estimar. Segundo Miraldo *et al.* [1], considerando que a variação entre *pixels* e rectas correspondentes é relativamente suave e com uma função  $\mathbf{f} : P^2 \rightarrow L^3$ , que faz o mapeamento de um ponto no plano da imagem para uma recta no espaço 3D, é possível reduzir a quantidade de cálculos a efectuar, assim como reduzir erros de ruído e cingir o número de parâmetros desconhecidos.

No modelo de imagem generalizado, apesar de vários pontos na imagem poderem corresponder ao mesmo raio no mundo, diferentes raios não podem ser mapeados no mesmo ponto da imagem (Figura 7), ou seja, o sistema  $\mathbf{l} = \mathbf{f}(\mathbf{x})$  assume a forma não injectiva. Por este motivo, o modelo de imagem generalizado apenas pode ser usado no mapeamento de pontos da imagem em rectas 3D, pois uma recta 3D pode corresponder a mais do que um ponto no plano da imagem.



**Figura 7** - Figura exemplificativa da função  $I=f(x)$  e da sua característica não injectiva. Na imagem à esquerda é representado um esquema de mapeamento duplo do mesmo *pixel* em duas rectas que não pode ocorrer ( $m_{12} \rightarrow l_{12}$  e  $m_{13} \rightarrow l_{13}$ ). Na imagem à direita está representado o esquema em que dois *pixels* dão origem a uma só recta ( $m_{21} \rightarrow l_{21}$  e  $m_{22} \rightarrow l_{21}$ ).

A escolha da representação de rectas 3D proveniente da referida função de mapeamento deve ser cuidada e por isso, a escolha das coordenadas de *Plücker* (Secção 3.2) assume aqui elevada importância. Uma recta 3D tem quatro graus de liberdade, no entanto, nenhuma das representações de apenas quatro variáveis contém toda a informação da recta. Por outro lado, as coordenadas de *Plücker* possuem seis elementos para representar os quatro graus de liberdade mas, têm a desvantagem da restrição de ortogonalidade entre os vectores direcção e momento. Segundo Bartoli e Sturm [14], e atendendo ao algoritmo proposto pelos mesmos autores, é benéfico utilizar o conceito de ortogonalidade e separar os vectores direcção e momento aquando da estimação. Assim,

a estimação da recta  $\hat{\mathbb{R}} = (\hat{\mathbf{d}}, \hat{\mathbf{m}})$  pode ser verificada ignorando a condição de ortogonalidade. Ao procurar a matriz de rotação mais próxima de,

$$\begin{pmatrix} \bar{\mathbf{d}} & \bar{\mathbf{m}} & \bar{\mathbf{d}} \times \bar{\mathbf{m}} \end{pmatrix}, \text{ tal que } \bar{\mathbf{x}} = \frac{\mathbf{x}}{\|\mathbf{x}\|} \quad (19)$$

ou seja, encontrando os vectores ortogonais mais próximos, é possível encontrar uma solução onde a direcção e o momento sejam ortogonais.

Tal como foi indicado anteriormente, em [1] foram utilizadas *RBF* na interpolação (Secção 3.1).

$$s(\mathbf{x}) = a_0 + \mathbf{a}_x^T \mathbf{x} + \sum_{i=1}^P \omega_i \phi(\|\mathbf{x} - \mathbf{c}_i\|) \quad (20)$$

onde  $a_0, \mathbf{a}_x$  e  $\omega_i$  são as incógnitas da interpolação,  $\phi$  é a RBF,  $\mathbf{c}_i \in \mathbb{R}^2$  e  $\mathbf{x} \in \mathbb{R}^2$ .

Separando as incógnitas na equação anterior e representando em forma matricial temos,

$$s(\mathbf{x}) = \begin{pmatrix} \phi(\mathbf{x}) & \mathbf{p}(\mathbf{x}) \end{pmatrix} \begin{pmatrix} \mathbf{w} \\ \mathbf{a} \end{pmatrix} \quad (21)$$

onde  $\mathbf{a} = (a_0, \mathbf{a}_x)$ ,  $\mathbf{w} = (\omega_1, \dots, \omega_p)$ ,  $\phi_i(\mathbf{x}) = \phi(\|\mathbf{x} - \mathbf{c}_i\|)$  e  $\mathbf{p}(\mathbf{x}) = \begin{pmatrix} 1 & x_1 & x_2 \end{pmatrix}$ .

No entanto, precisamos de seis interpoladores para estimar os seis elementos de  $\hat{\mathbb{R}}$ . Então, podemos escrever o vector  $\mathbf{s}(\mathbf{x})$  da seguinte forma

$$\mathbf{s}(\mathbf{x}) = \begin{pmatrix} s_1(\mathbf{x}) & s_2(\mathbf{x}) & s_3(\mathbf{x}) & s_4(\mathbf{x}) & s_5(\mathbf{x}) & s_6(\mathbf{x}) \end{pmatrix} \quad (22)$$

combinando a informação da equação (21) com a equação (22),

$$\mathbf{s}(\mathbf{x}) = \begin{pmatrix} \phi(\mathbf{x}) & \mathbf{p}(\mathbf{x}) \end{pmatrix} \mathbf{H}_{wa} \quad (23)$$

Em que,

$$\mathbf{H}_{wa} = \begin{pmatrix} \mathbf{h}_{wa}^{(1)} & \mathbf{h}_{wa}^{(2)} & \mathbf{h}_{wa}^{(3)} & \mathbf{h}_{wa}^{(4)} & \mathbf{h}_{wa}^{(5)} & \mathbf{h}_{wa}^{(6)} \end{pmatrix} \quad (24)$$

Então a estimação do vector de *Plücker* correspondente à recta no espaço que queremos saber é,

$$\hat{\mathbf{l}}\mathbb{R} = \mathbf{s}^T, \quad \mathbf{s}(\mathbf{x}) \in \mathbb{R}^{1 \times 6} \quad (25)$$

Finalmente, segundo as equações (23) e (25)

$$\hat{\mathbf{l}}^T \mathbb{R} = \begin{pmatrix} \phi(\mathbf{x}) & \mathbf{p}(\mathbf{x}) \end{pmatrix} \mathbf{H}_{wa} \quad (26)$$

Uma vez calculada a matriz  $\mathbf{H}_{wa}$  para um sistema de imagem, esta pode ser usada para calcular as rectas correspondentes a um determinado ponto  $\mathbf{x}$  na imagem. Os valores de  $\mathbf{H}_{wa}$  dependem do conjunto de pontos  $\{\mathbf{c}_i\}$  utilizados na RBF (Secção 3.1). Quando consideramos o mesmo sistema de imagem e alteramos o conjunto  $\{\mathbf{c}_i\}$  teremos necessariamente uma matriz  $\mathbf{H}_{wa}$  diferente. Esta matriz depende de  $\{\mathbf{c}_i\}$ , daí a denominação “pontos de controlo”. No entanto, se utilizarmos os mesmos pontos de controlo para dois sistemas de imagem diferentes, apenas a matriz  $\mathbf{H}_{wa}$  será diferente. Podemos então dizer que a matriz  $\mathbf{H}_{wa}$  é a matriz do sistema de imagem ou matriz da câmara.

Concluimos que, para um sistema de imagem genérico para o qual é previamente definido um conjunto de pontos de controlo  $\{\mathbf{c}_i\}$ , é possível obter uma função de mapeamento  $\mathbf{s} : \mathbb{R}^2 \rightarrow \mathbb{R}^6$  através da estimação de uma matriz de câmara  $\mathbf{H}_{wa}$  pertencente a  $\mathbb{R}^{(P+3) \times 6}$  constituída por  $6P+18$  incógnitas que necessariamente têm de ser estimadas para ser possível resolver o sistema.

### 3.4. Calibração

Tal como foi referido anteriormente (Secção 2.2), ambos os processos de calibração sugeridos por Grossberg e Nayar [2, 3] e Sturm e Ramalingam [4, 5], são computacionalmente muito pesados por requererem a associação de um *pixel* da imagem a uma recta para todos os *pixels* (sistemas discretos). Além disso, para o cálculo de cada recta são necessários dois ou mais pontos 3D. Assim, a simplificação e racionalização de dados é pertinente e importante para obter um método exequível. A parametrização da representação do modelo de imagem genérico [1] descrito anteriormente, permite simplificar os cálculos e reduzir o número de variáveis a calcular. Além disso, em vez de dois ou mais pontos 3D, o método proposto [1] apenas necessita de um ponto 3D para cada ponto da imagem.

Sabemos da secção anterior (Secção 3.3) que a matriz do sistema de imagem ou  $\mathbf{H}_{wa}$  é constituída por um grupo de incógnitas. Os pontos de controlo  $\{\mathbf{c}_i\}$ , que pertencem ao segundo grupo de incógnitas, são definidos inicialmente através de um processo de *clustering*. Este é baseado na selecção de conjuntos de pontos da imagem próximos entre si. Em seguida, é feita a selecção de um ponto médio que fica identificado como ponto de controlo. Aplicando à equação  $\mathbf{Q}(\mathbf{p})$ , (equação (18))

$$(\mathbf{Q}(\mathbf{p})\hat{\mathbf{I}}\mathbb{R})^T = \hat{\mathbf{I}}\mathbb{R}^T \mathbf{Q}(\mathbf{p})^T = 0 \quad (27)$$

substituindo  $\hat{\mathbf{I}}\mathbb{R}$  pela equação (26),

$$\mathbf{r}(\mathbf{x})\mathbf{H}_{wa}\mathbf{Q}(\mathbf{p})^T = 0 \quad (28)$$

em que  $\mathbf{r}(\mathbf{x})$  é dado por

$$\left( \begin{array}{cc} \phi(\mathbf{x})^T & \mathbf{p}(\mathbf{x})^T \end{array} \right) = \mathbf{r}(\mathbf{x}), \quad \mathbf{r}(\mathbf{x}) \in \mathbb{R}^{1 \times (P+3)} \quad (29)$$

Utilizando o produto de Kronecker (Secção 1.2) podemos isolar as incógnitas. A matriz do produto de Kronecker pertence a  $\mathbb{R}^{4 \times (6P+18)}$ ,  $vec(\mathbf{H}_{wa}) \in \mathbb{R}^{(6P+18)}$  e resulta da vectorização de  $\mathbf{H}_{wa}$  numa coluna.

Isolando  $\mathbf{H}_{wa}$  obtemos a seguinte equação

$$\left[ \mathbf{Q}(\mathbf{p}) \otimes \mathbf{r}(\mathbf{x}) \right] vec(\mathbf{H}_{wa}) = 0 \quad (30)$$

Para  $N$  correspondências entre pontos da imagem e pontos do mundo  $\{\mathbf{x}_i \rightarrow \mathbf{p}_i\}$ , é possível construir um sistema através da equação (27).

$$\begin{cases} \mathbf{Q}(\mathbf{p}_1) \hat{\mathbf{l}}_1 \mathbb{R} = 0 \\ \mathbf{Q}(\mathbf{p}_2) \hat{\mathbf{l}}_2 \mathbb{R} = 0 \\ \vdots \\ \mathbf{Q}(\mathbf{p}_N) \hat{\mathbf{l}}_N \mathbb{R} = 0 \end{cases} \quad (31)$$

O sistema pode ser transformado numa matriz denominada  $\mathbf{M}$  através das equações (28) e (30),

$$\mathbf{M} = \begin{pmatrix} \mathbf{Q}(\mathbf{p}_1) \otimes \mathbf{r}(\mathbf{x}_1) \\ \mathbf{Q}(\mathbf{p}_2) \otimes \mathbf{r}(\mathbf{x}_2) \\ \vdots \\ \mathbf{Q}(\mathbf{p}_N) \otimes \mathbf{r}(\mathbf{x}_N) \\ \mathbf{D} \end{pmatrix} \quad (32)$$

onde  $\mathbf{M} \in \mathbb{R}^{(4N+18) \times (6P+18)}$  e a matriz  $\mathbf{D} \in \mathbb{R}^{18 \times (6P+18)}$  de acordo com Miraldo *et al.* [12].

Concluimos que a calibração passa então pela estimação da matriz  $\mathbf{H}_{wa}$  na forma de vector tal que

$$\mathbf{M}vec(\mathbf{H}_{wa}) = 0 \quad (33)$$

o que significa que  $vec(\mathbf{H}_{wa})$  pertence ao espaço nulo de  $\mathbf{M}$ .

Para assegurar uma solução única para o sistema, temos de garantir que a característica de  $\mathbf{M}$  é  $6P + 17$  uma vez que  $vec(\mathbf{H}_{wa})$  pertence a  $\mathbb{R}^{(6P+18)}$ , ou seja, o espaço nulo de  $\mathbf{M}$  será igual a 1. Com vista a esse objectivo, consideramos que qualquer elemento do espaço nulo de  $\mathbf{M}$ , excepto  $vec(\mathbf{H}_{wa}) = 0$  é solução a menos de um factor de escala (equação (27)). Ainda segundo Miraldo *et al.* [1, 12], pode ser provado que é possível obter  $\mathbf{M}$  com característica  $6P+17$ , definindo  $N = 2P$ , para uma matriz  $\mathbf{D} \in \mathbb{R}^{18 \times (6P+18)}$ , que é obtida a partir das equações (14). É necessária a imposição de algumas restrições na relação entre pontos de controlo  $\{\mathbf{c}_i\}$  e pontos da imagem  $\{\mathbf{x}_i\}$ . É também necessária uma escolha regrada da quantidade de pontos de controlo de acordo com a RBF escolhida e com os pontos assinalados na imagem. Foram sugeridas três tipos de *radial basis functions* que podem ser utilizadas (Secção 3.1.1).

No entanto, em termos práticos, devido essencialmente ao ruído da imagem e tendo em consideração a quantidade de pontos que é possível obter através do método proposto [1], não é relevante a imposição de todas estas regras. Considerando que o número de pontos obtidos no *dataset* é várias vezes superior ao número de pontos de controlo, temos a garantia que  $\mathbf{M}$  tem a característica máxima [12]. Ou seja, para  $\mathbf{x}_i \gg \mathbf{c}_i$  podemos assumir que  $\mathbf{M}$  tem característica  $6P + 18$ . Fazendo a decomposição SVD (Secção 1.2), vamos obter o vector que minimiza a solução da equação (33).

$\mathbf{M} = \mathbf{U}\mathbf{D}\mathbf{V}^T$  e  $\mathbf{x} = \mathbf{V}(:,n)$  em que  $n$  é o número de colunas da matriz  $\mathbf{M}$ .



# CAPÍTULO 4 - UM PONTO NO MUNDO, VÁRIOS PONTOS NA IMAGEM

## 4.1. O problema das multi correspondências em modelos de imagem

Quando o sistema de imagem é composto por várias perspectivas, é frequente que um ponto no mundo seja visto em vários pontos da imagem. Num sistema ideal as duas rectas cruzam-se nesse ponto, no entanto, em sistemas com ruído esse facto não acontece (Figura 8). Um ponto incidente com uma recta no mundo satisfaz a seguinte equação,

$$\mathbf{p} \times \mathbf{d} - \mathbf{m} = 0 \quad (35)$$

em que  $\mathbf{p}$  é um ponto em  $P^3$ ,  $\mathbf{d}$  e  $\mathbf{m}$  são os vectores direcção e momento da recta. A equação pode ser transformada para,

$$[\mathbf{d}]_x \mathbf{p} + \mathbf{m} = 0 \quad (36)$$

onde  $\mathbf{p}$  pertence a  $\mathbb{R}^3$  e  $[\mathbf{a}]_x$  é uma matriz que lineariza o produto externo como  $[\mathbf{a}]_x \mathbf{b} = \mathbf{a} \times \mathbf{b}$ .

A equação pode ser representada em forma matricial como,

$$\begin{bmatrix} [\mathbf{d}]_x & \mathbf{m} \end{bmatrix} \begin{bmatrix} \mathbf{p} \\ 1 \end{bmatrix} = 0 \quad (37)$$

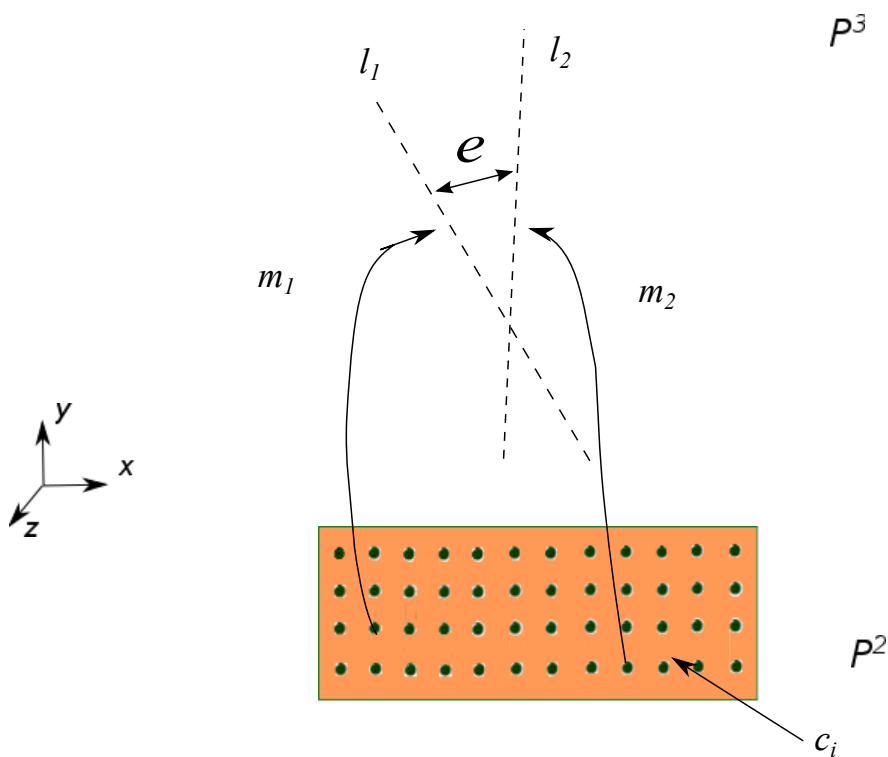
Para várias rectas incidentes em  $\mathbf{p}$ , podemos escrever,

$$\mathbf{B} \begin{bmatrix} \mathbf{p} & 1 \end{bmatrix}^T = 0 \quad (38)$$

em que,

$$\mathbf{B} = \begin{bmatrix} [\mathbf{d}_1]_x & \mathbf{m}_2 \\ [\mathbf{d}_2]_x & \mathbf{m}_2 \\ \vdots & \vdots \\ [\mathbf{d}_n]_x & \mathbf{m}_n \end{bmatrix} \quad (39)$$

Para um sistema sem ruído, a característica de  $\mathbf{B}$  é três. O espaço nulo de  $\mathbf{B}$  vai conter as soluções para o sistema. Quando existe ruído (Figura 8), seleccionamos o  $\mathbf{p}$  que minimiza o sistema da equação (38), fazendo a SVD( $\mathbf{B}$ ) e seleccionando a última coluna da matriz  $\mathbf{V}$  (Secção 1.2).

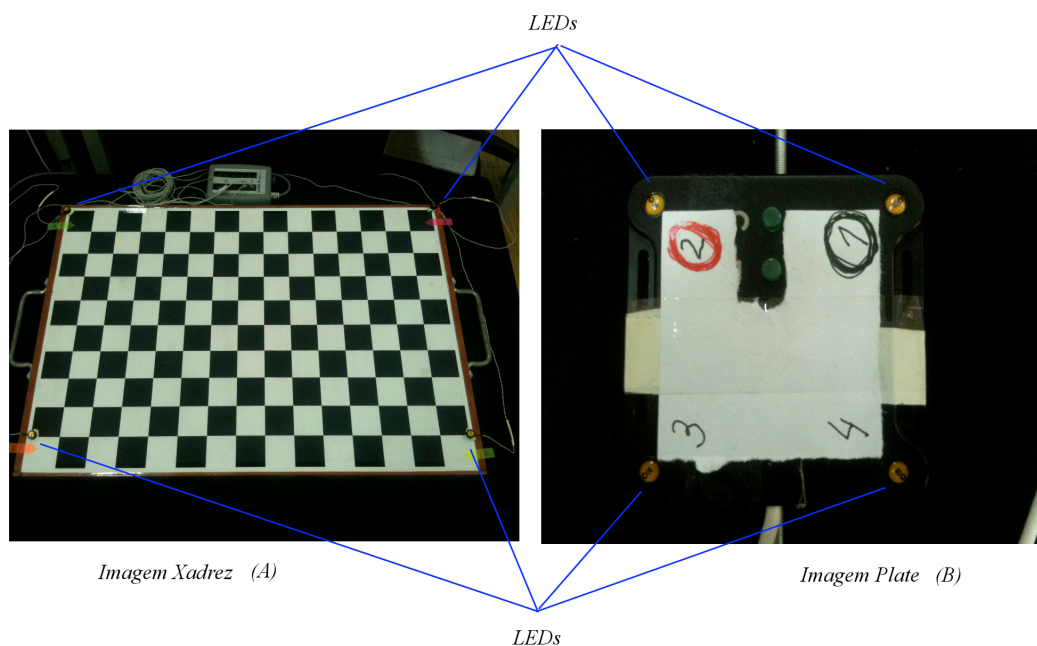


**Figura 8** - Exemplo de um caso em que duas rectas correspondentes ao mesmo ponto 3D são mapeadas por dois *pixels* diferentes. Para o caso ideal as rectas intersectam-se no ponto 3D. Quando existe ruído as rectas estão separadas por uma distância  $e$ .

# CAPÍTULO 5 - IMPLEMENTAÇÃO PRÁTICA

## 5.1. Calibração

Para a aplicação prática, utilizámos três sistemas de imagem diferentes (Secção 2.1), que calibrámos de acordo com o método proposto por Miraldo *et al.* [1]. No primeiro caso, testámos um sistema projectivo comum, no segundo um sistema catadioptrico com um espelho esférico e no terceiro um sistema catadióptrico com dois espelhos planos. Para ter a correspondência 2D para 3D de modo a que pudéssemos testar a precisão do sistema, utilizámos um sistema de medição por infravermelhos (IR) (Figura 10). Colocámos quatro LEDs num tabuleiro de xadrez (Figura 9 (A)) para a parte de calibração e utilizámos uma placa fornecida com o dispositivo (*plate*) (Figura 9 (B)) de medição, para efectuar os testes comparativos.



**Figura 9** - Imagens do Xadrez e do *plate* utilizados na recolha de datasets, de treino (A) e de resultados (B).

O sistema de medição *Optotraker* tem uma precisão de 0.1mm e uma resolução de 0.01mm [15]. Os quatro LEDs foram colocados nos cantos do xadrez. Desta forma, sabendo o tamanho do xadrez e das suas quadrículas, é possível saber aproximadamente a posição dos vértices correspondentes a todas as quadrículas  $\{\mathbf{p}_i\}$ . Os respectivos vértices vistos na imagem  $\{\mathbf{x}_i\}$  serão marcados pelo utilizador no processo de calibração e será esta a associação entre pontos  $\{\mathbf{x}_i \rightarrow \mathbf{p}_i\}$ . Com a relação entre pontos na imagem e pontos no mundo, estimamos a função de interpolação  $s$ . Tal como foi explicado anteriormente (Secção (3.4)),  $P$  deve ser escolhido tendo em consideração o número de pontos do *dataset*. Queremos que a condição  $N \gg P$  seja cumprida, por isso o número de pontos do *dataset* deve ser muito maior que o número de pontos de controlo. Para obter uma solução para a equação (33), podemos usar o método dos mínimos quadrados (Secção 1.2). Após a recolha de pontos e imagens com o xadrez procedemos à medição com o *plate*. Nesta fase é gerado um novo sistema de correspondências entre pontos da imagem e pontos do mundo, que servirá para medir a precisão da calibração efectuada. Assim, temos um novo conjunto de dados  $\{\mathbf{y}_i \rightarrow \mathbf{w}_i\}$ , onde  $\mathbf{y}_i$  são pontos da imagem e  $\mathbf{w}_i$  são pontos do mundo. O erro será a distância entre o ponto no mundo  $\mathbf{p}_i$  para a recta gerada  $\hat{\mathbf{l}}_i$ , onde  $\hat{\mathbf{l}}_i = \mathbf{s}(\mathbf{y}_i)$ .



**Figura 10** - Sistema de medição IR usado para fazer as medições *Optotraker* [15].

### **5.1.1. Aquisição do dataset**

Como passo inicial é necessário colocar os LEDs no xadrez, nos cantos do tabuleiro (Figura 9 (A)). É também importante saber as medidas do xadrez (número de quadrados e medida dos mesmos). De seguida, é feita a montagem do sistema com a câmara e o sistema de medição. O sistema de medição requer software/hardware muito específico. Apesar de a resolução não interferir com a calibração directamente, acaba por ter o seu peso devido aos erros de marcação por parte do utilizador. Antes de iniciar a aquisição de imagens para a fase de calibração, é de extrema importância fixar a câmara para evitar deslocações que podem influenciar a calibração. Seguidamente, são adquiridas duas sequências de imagens. Em primeiro lugar, variamos a posição 3D do xadrez, com o cuidado de verificar que todos os LEDs se encontram detectáveis pelo sistema de IR *Optotraker*. As imagens são adquiridas quando o utilizador considera que o xadrez está na posição pretendida (Figuras 11 e 12). É importante que nesta sequência de imagens seja abrangida grande parte da área de imagem com diferentes profundidades. Após este processo é feita a aquisição da segunda sequência de imagens com o *plate*. Esta tem a função de testar a precisão da calibração efectuada. As imagens são adquiridas numa sequência contínua (7,5 frames por segundo), onde o utilizador apenas assinala o início e o fim. Em cada uma das fases são registadas num ficheiro as coordenadas 3D associadas aos LEDs em cada imagem, dados que serão usados para a calibração e mais tarde para comparação dos resultados.

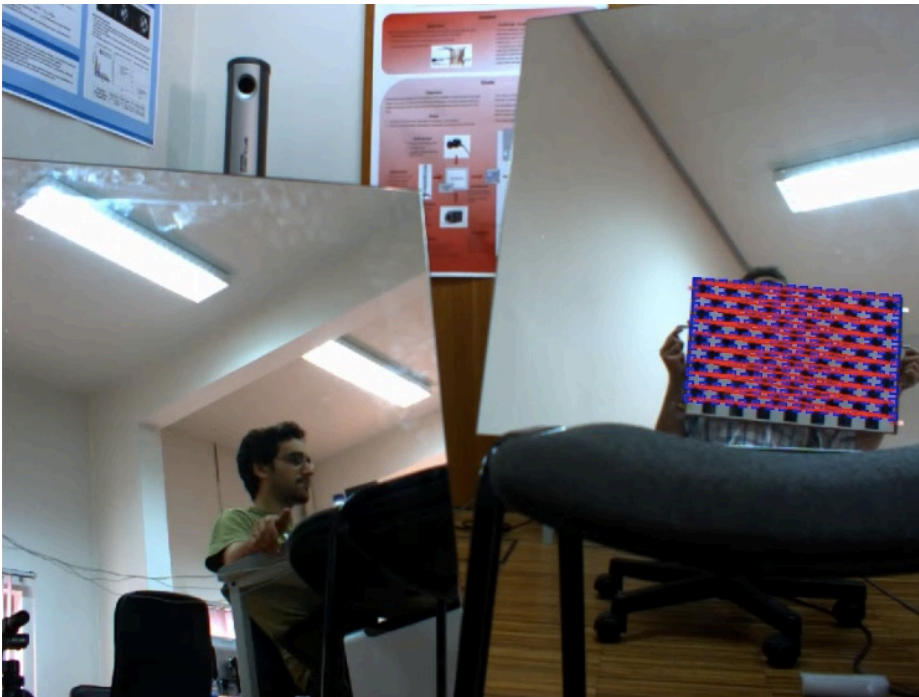
### **5.1.2. Calibração**

Quando estamos na posse de todos os dados podemos proceder à calibração. É um processo moroso e que requer alguma sensibilidade por parte do utilizador, especialmente na selecção dos pontos. Nem sempre é possível ter a qualidade de imagem necessária para seleccionar exactamente o *pixel* pretendido. Inicialmente, é necessário definir o número de pontos de controlo e marcar todos os vértices das quadrículas do xadrez em todas as imagens adquiridas. O xadrez tinha 120 pontos por imagem (Figura 9 (A)), que correspondem a todos os pontos do xadrez que estão situados entre os 4 LEDs utilizados. A marcação pelo utilizador dará origem ao conjunto de pontos da imagem  $\{\mathbf{x}_i\}$ . Em seguida, são efectuados os cálculos para a calibração onde se obtém a matriz  $\mathbf{M}$  e todos os dados necessários para calibrar o referido sistema de imagem.

Numa segunda fase é feita a marcação dos pontos  $\{\mathbf{y}_i\}$ . Este procedimento tem como objectivo testar a qualidade dos resultados da calibração. A marcação é feita de forma similiar à anterior, no entanto a quantidade de pontos é muito menor, apenas quatro por imagem (Figura 9 (B)).

### 5.1.3. Resultados

Por último, é feita a comparação de resultados entre os dados adquiridos pelo sistema de medição *Optotraker* (Figura 10) e os resultados gerados pelas calibrações efectuadas utilizando as diferentes *RBF* (Secção 3.1.1). Para cada sistema de imagem são calculados os erros referentes aos quatro

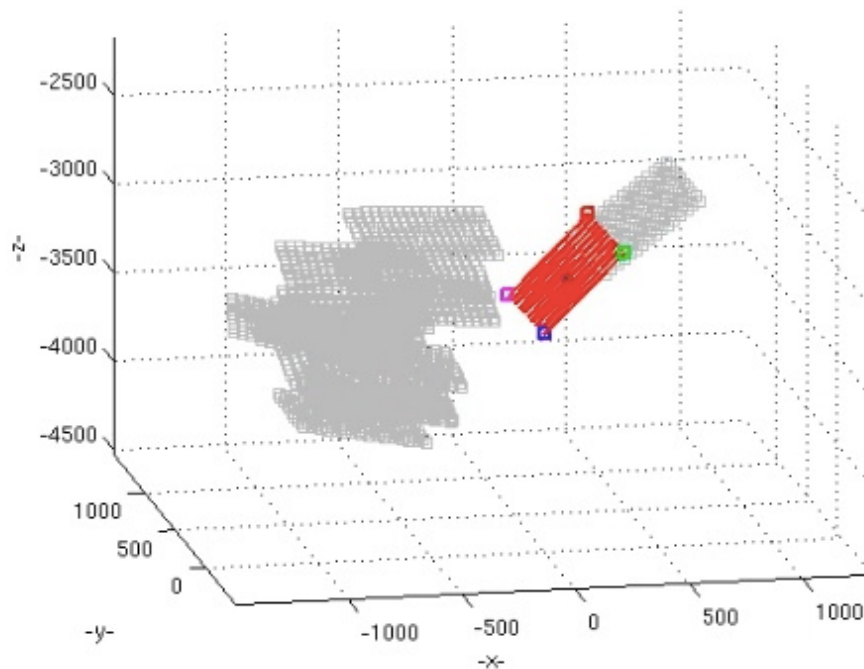


**Figura 11** - Imagem da fase de calibração do sistema catadioptrico com dois espelhos planos, onde é possível observar a disposição do sistema com o *Optotraker* ao fundo e a câmara direccionada para os dois espelhos. É também possível observar o xadrez utilizado.

LEDs para cada tipo de *RBF*. Este erro é dado pela distância do ponto no mundo  $\mathbf{p}_i$  à linha gerada na calibração,  $\hat{\mathbf{l}}_i$  onde  $\hat{\mathbf{l}}_i = \mathbf{s}(\mathbf{y}_i)$ . Um conjunto de  $\{\mathbf{y}_i\}$  e linhas correspondentes  $\{\hat{\mathbf{l}}_i\}$  será apresentado em imagens futuras (Secção 6). A distância entre uma linha, em coordenadas de *Plücker*, e um ponto do mundo é dada por,

$$\varepsilon_i = \left\| \begin{pmatrix} [\mathbf{w}_i]_{\times} & -I \end{pmatrix} \hat{\mathbf{l}}_i \right\| \quad (34)$$

onde  $\bar{\mathbf{l}}_i = \hat{\mathbf{l}}_i / \|\mathbf{d}_i\|$ ,  $\mathbf{d}$  corresponde ao vector direcção da recta e  $\mathbf{w}$  é um ponto 3D.



**Figura 12** - Imagem do referencial utilizado com origem no *Optotraker*. A vermelho estão os pontos marcados pelo utilizador nessa interacção e a cinzento as interacções anteriores. Os quatro LEDs estão também assinalados.

## 5.2. Algoritmo Computacional

1 - Definição das constantes base da calibração:

- Medidas do xadrez:
  - Quantidade de quadrículas;
  - Medida do lado de uma quadrícula;
  - Posição dos LEDs no xadrez.

2 - Marcação dos pontos nas diversas imagens:

- Marcação pelo utilizador e confirmação de dados;
- Correspondência dos pontos marcados com os pontos adquiridos pelo sistema de IR;

3 - Calibração:

- Distribuição dos pontos de controlo (*cluster* de  $\{\mathbf{x}_i\}$ );
- Construção da matriz  $\mathbf{M}$  (equação (33));

\_ Cálculo matriz  $\mathbf{H}_{wa}$  para as diversas RBF.

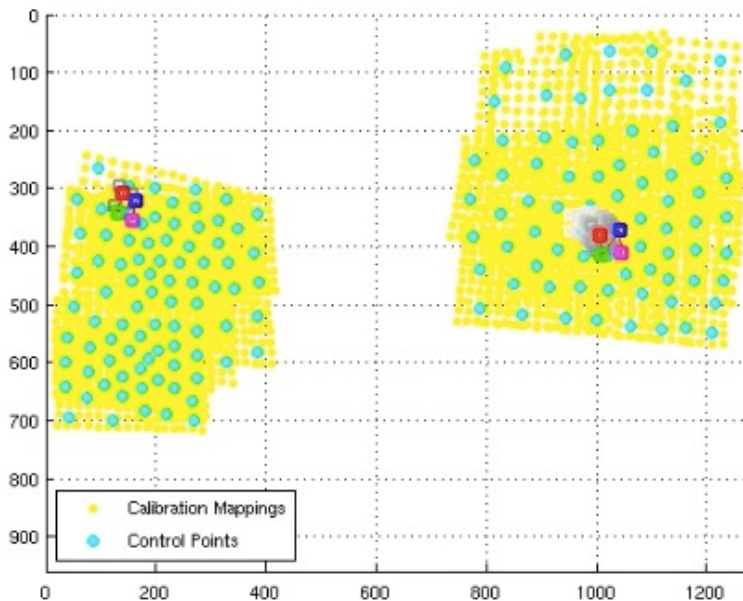
#### 4 - Comparação de resultados:

- Filtragem de pontos sem correspondência (pontos em que o sistema de medição não detectou pelo menos três LEDs);
- Cálculo das distâncias de erro (Secção 5.1);
- Gerar resultados:
  - Criação de sequência de imagens;
  - Cálculo da média e desvio padrão.



## CAPÍTULO 6 - RESULTADOS

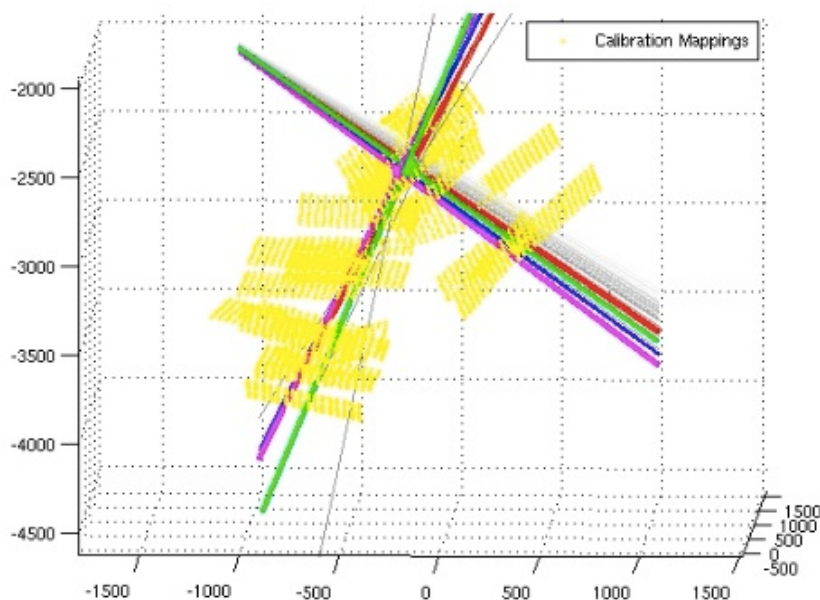
Os testes práticos que efectuámos conduziram a um longo percurso de aprendizagem e experimentação onde, tentámos aperfeiçoar o método proposto [1] para obter os melhores resultados possíveis.



**Figura 13** - Imagem da fase de testes. A amarelo estão os pontos do dataset e a azul os pontos de controlo. Estão também representados os pontos correspondentes aos quatro LEDs do *plate* que originam as rectas da Figura 14. É possível observar as diferentes distribuições de pontos de controlo.

Um dos primeiros problemas com que nos deparámos foi a existência de um desfasamento entre os dados obtidos pelo sistema de medição *Optotraker* e as imagens que estávamos a adquirir. Considerando que a taxa de aquisição era 7,5 frames por segundo, um erro deste género é enorme. De notar que, apenas na segunda sequência de imagens é que este problema é crítico pois, como foi dito anteriormente (Secção 5.1.1), esta é obtida continuamente (7,5 frames por segundo). A limitação das frames originou outro problema. Quando existiam movimentos mais rápidos com o *plate*, o problema do desfasamento acentuava-se e gerava erros de medição. Ainda relacionado com os erros de medição, deparamo-nos com a situação em que o sistema de medição *Optotraker* detectava menos de três LEDs. Quando esta situação ocorre o sistema atribui um valor numérico elevado à distância. Se este valor não for eliminado dos cálculos, vai provocar grandes desvios na comparação dos valores das rectas.

A simplicidade do modelo projectivo associado à facilidade de marcação dos pontos, resultou numa combinação com excelentes resultados e diferenças mínimas, nas correspondências entre o valor calculado e o valor real. No caso do sistema catadioptrico com uma esfera espelhada, os resultados foram bastante positivos mas mais fracos que no caso projectivo. A principal razão para esta situação é o facto de a marcação dos pontos ter sido extremamente complicada de fazer. Tanto o



**Figura 14** - Imagem da fase de testes, onde são desenhadas as rectas obtidas na calibração correspondentes aos quatro pontos do plate da Figura 13.

xadrez como o *plate* ficam muito pequenos na imagem após a sua reflexão na superfície espelhada. No caso do *plate* houve situações em que os quatro pontos eram praticamente coincidentes. É importante destacar que apesar de todos estes problemas, os resultados adquiridos continuam a ter erros mínimos. Acreditamos que o sistema com a esfera podia ter melhores resultados se a imagem não fosse reduzida devido ao espelho. Finalmente, para o caso catadioptrico com dois espelhos planos, os resultados foram semelhantes ao sistema projectivo. Neste último caso, demonstrámos que o sistema funciona para múltiplas câmaras, pelo que é possível observar as diferentes perspectivas em simultâneo (Figura 14).

Foram testadas três *RBF* (Secção 3.1.1). Existe uma relação entre a área da imagem calibrada e o número de pontos de controlo. Caso este número seja muito elevado os resultados serão piores, para um número fixo de pontos de um *dataset*. O mesmo acontece para o caso em que o número de pontos de controlo é baixo. No primeiro caso, o número elevado de pontos de controlo leva à existência “sobreposições” na interpolação. Este facto leva ao cálculo de aproximações erradas. No caso de o número de pontos de controlo ser demasiado baixo, leva a que a complexidade do modelo paramétrico não acompanhe a complexidade do modelo que representa (Figura 13).

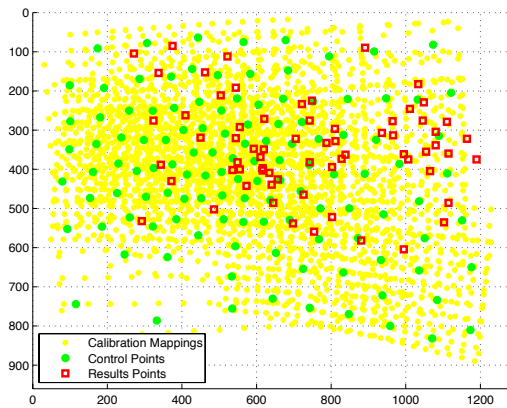
Após a análise de resultados, é fácil notar que em qualquer um dos modelos de imagem, a *RBF gaussiana* apresenta uma sensibilidade ao ruído extremamente elevada. O mínimo desvio provoca erros enormes na estimação das rectas. Podemos observar (Tabela 2) que a *RBF gaussiana* é a que apresenta erro com maior média e desvio padrão. Consideramos que na prática, a *RBF gaussiana* não possui grande espaço de melhoramento, excepto se for possível reduzir o ruído de forma considerável.

Em relação às *RBF multiquádricas* os resultados são substancialmente melhores. Mesmo quando ocorrem movimentos bruscos e o problema da *framerate* se torna evidente, este tipo de *RBF* continua a manter resultados bastante precisos. As médias são aproximadamente um centímetro (Tabela 2) para os casos projectivo e multiprojectivo. Já para o caso esférico, são um pouco superiores, cerca de dois centímetros, devido à questão do espelho reduzir a imagem.

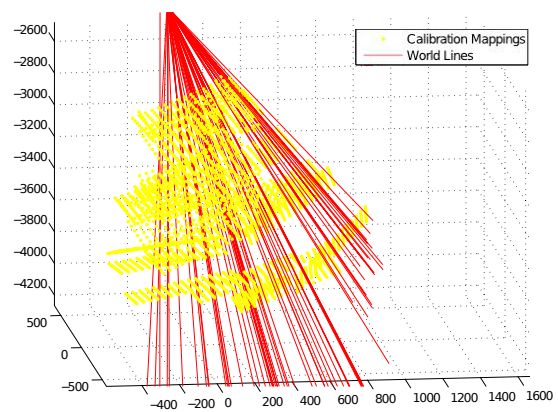
Decidimos também testar um novo tipo de *RBF*, as *thin-plate-splines*, que ao contrário das duas anteriores não têm parâmetro de forma (Tabela 1). Os resultados foram muito positivos. Mais estáveis e resistentes a ruído, com as *thin-plate-splines* conseguimos obter resultados melhores do que com as *multiquádricas*. A ausência de parâmetro de forma nas *thin-plate-splines* é um factor determinante para a qualidade dos resultados obtidos.

Apresentamos ainda algumas imagens dos modelos usados (Secção 2.1), com representações das rectas 3D (Figura 15b) 16b) e 17b)), dos pontos de controlo e *pixels* na imagem (Figura 15a) 16a) e 17a)). Mostramos também algumas imagens com os valores dos erros (Figura 15c) 16c) e 17c)). Na Tabela 2 expomos os resultados obtidos para a média e desvio padrão referentes aos três tipos de modelo de imagem e para as três *RBF* usadas.

Apresentamos ainda imagens do caso em que múltiplas perspectivas aparecem representadas (Figura 18a) 18b) e 18c)). São calculados os erros dos pontos estimados, da intersecção das duas rectas em relação ao ponto 3D dado pelo sistema de medição *Optotraker* e a média e desvio padrão são apresentados na Tabela 3. Finalmente, na Tabela 4 apresentamos dados relativos ao número de pontos utilizados (pontos de controlo, de calibração e de teste).



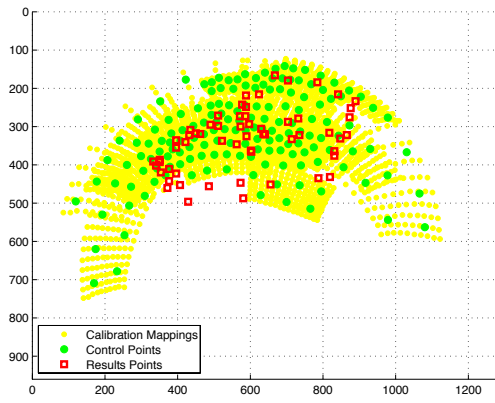
**Figura 15a)** - Imagem dos resultados obtidos para o modelo projectivo, onde os pontos amarelos correspondem ao conjunto  $\{x_i\}$ , os pontos verdes são pontos de controlo usados na calibração e os quadrados vermelhos  $\{y_i\}$  são parte do conjunto  $\{y_i \rightarrow \hat{l}_i\}$  gerado com *thin-plate-splines*.



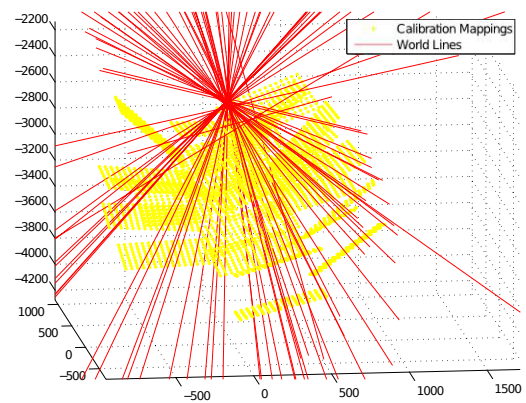
**Figura 15b)** - Imagem dos resultados obtidos para o modelo projectivo, onde os pontos amarelos correspondem ao conjunto  $\{y_i\}$  e as rectas vermelhas  $\{\hat{l}_i\}$  são parte do conjunto  $\{y_i \rightarrow \hat{l}_i\}$  gerado com *thin-plate-splines*.



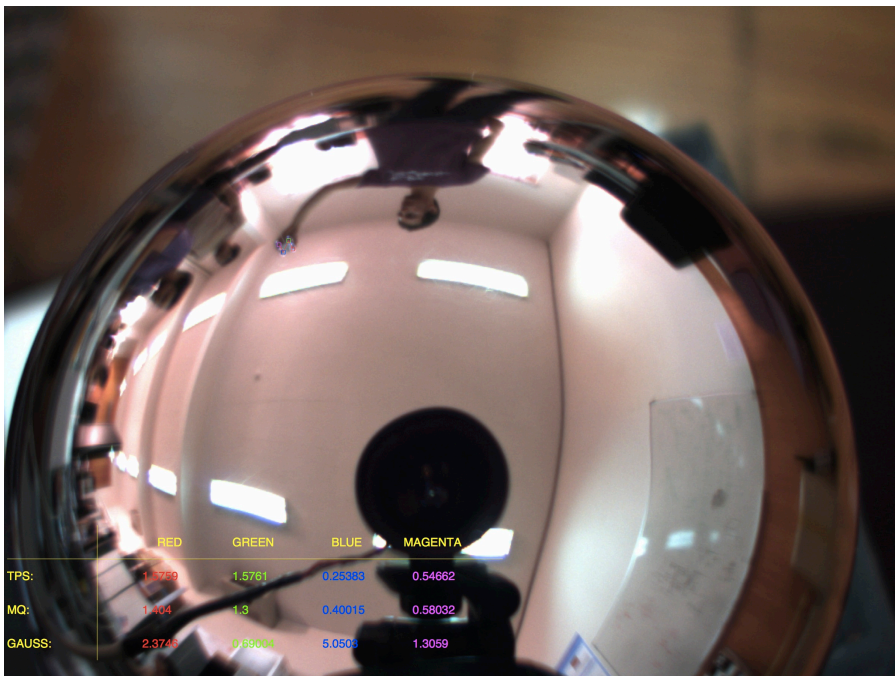
**Figura 15c)** - Imagem dos resultados obtidos para o modelo projectivo, com apresentação de resultados para os três tipos de *RBF*.



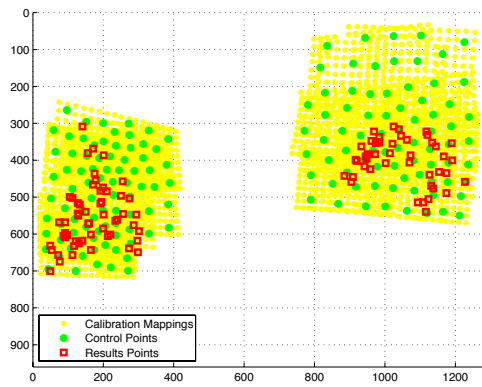
**Figura 16a)** - Imagem dos resultados obtidos para o modelo catadioptrico com espelho esférico, onde os pontos amarelos correspondem ao conjunto  $\{x_i\}$ , os pontos verdes são pontos de controlo usados na calibração e os quadrados vermelhos  $\{y_i\}$  são parte do conjunto  $\{y_i \rightarrow \hat{I}_i\}$  gerado com *thin-plate-splines*.



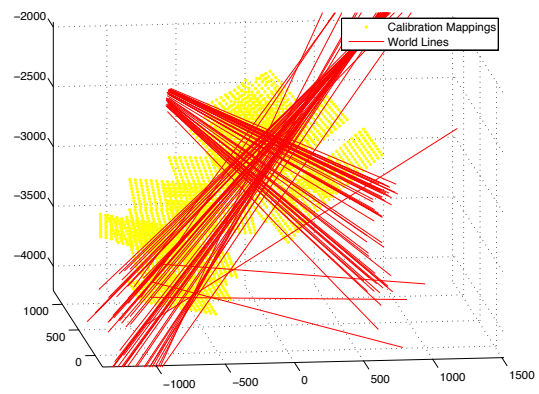
**Figura 16b)** - Imagem dos resultados obtidos para o modelo catadioptrico com espelho esférico, onde os pontos amarelos correspondem ao conjunto  $\{y_i\}$  e as rectas vermelhas  $\{\hat{I}_i\}$  são parte do conjunto  $\{y_i \rightarrow \hat{I}_i\}$  gerado com *thin-plate-splines*.



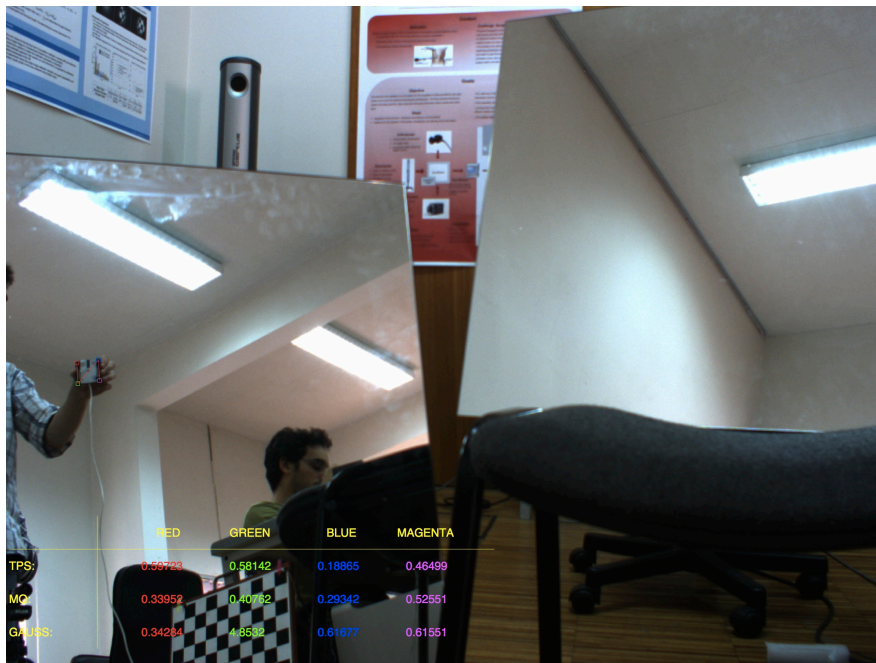
**Figura 16c)** - Imagem dos resultados obtidos para o modelo catadioptrico com espelho esférico, com apresentação de resultados para os três tipos de *RBF*.



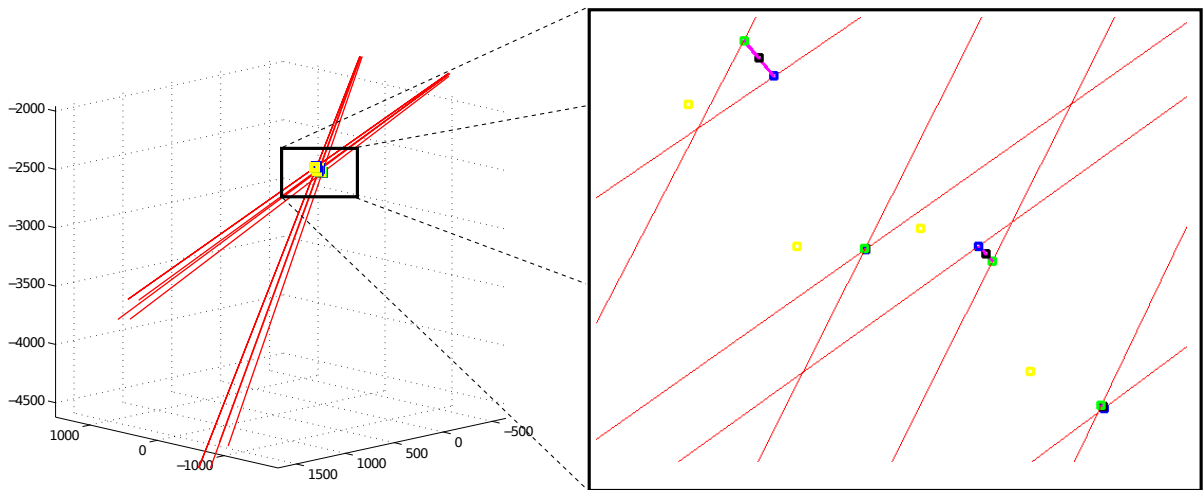
**Figura 17a)** - Imagem dos resultados obtidos para o modelo catadioptrico com dois espelhos planos, onde os pontos amarelos correspondem ao conjunto  $\{x_i\}$ , os pontos verdes são pontos de controlo usados na calibração e os quadrados vermelhos  $\{y_i\}$  são parte do conjunto  $\{y_i \rightarrow \hat{l}_i\}$  gerado com *thin-plate-splines*.



**Figura 17b)** - Imagem dos resultados obtidos para o modelo catadioptrico com dois espelhos planos, onde os pontos amarelos correspondem ao conjunto  $\{y_i\}$  e as rectas vermelhas  $\{\hat{l}_i\}$  são parte do conjunto  $\{y_i \rightarrow \hat{l}_i\}$  gerado com *thin-plate-splines*.



**Figura 17c)** - Imagem dos resultados obtidos para o modelo catadioptrico com dois espelhos planos, com apresentação de resultados para os três tipos de *RBF*.



**Figura 18a)** - Imagem dos resultados obtidos para o modelo catadioptrico com dois espelhos planos, num dado momento em que existem dois viewpoints. As rectas vermelhas  $\{\hat{l}_i\}$  são parte do conjunto  $\{y_i \rightarrow \hat{l}_i\}$  gerado com *thin-plate-splines*. Os pontos da zona de intersecção estão explicados na Figura 14b).

**Figura 18b)** - Imagem ampliada dos resultados obtidos para o modelo catadioptrico com dois espelhos planos, num dado momento em que existem dois viewpoints. As rectas vermelhas  $\{\hat{l}_i\}$  são parte do conjunto  $\{y_i \rightarrow \hat{l}_i\}$  gerado com *thin-plate-splines*. Tanto os quadrados azuis como os verdes são pontos de incidência das respectivas rectas. Os quadrados amarelos são pontos 3D dados pelo sistema de medição *Optotraker* e os quadrados pretos são a estimativa da intersecção das duas rectas.



**Figura 18c)** - Imagem dos resultados obtidos para o modelo catadioptrico com dois espelhos planos, num dado momento em que existem dois viewpoints. Os resultados apresentados são referentes aos três tipos de *RBF*.

Modelo de Imagem	RBF	Média	Desvio Padrão
Projectivo	<i>Gaussiana</i>	1.7009 (cm)	4.6529 (cm)
	<i>Multiquadrica</i>	0.6408 (cm)	0.7678 (cm)
	<i>Thin-Plate-Splines</i>	0.5883 (cm)	0.4915 (cm)
Catadioptrico com Espelho Esférico	<i>Gaussiana</i>	4.8993 (cm)	10.6219 (cm)
	<i>Multiquadrica</i>	2.4405 (cm)	3.5738 (cm)
	<i>Thin-Plate-Splines</i>	2.6508 (cm)	4.3103 (cm)
Catadioptrico com dois Espelhos Planos	<i>Gaussiana</i>	2.2523 (cm)	6.2586 (cm)
	<i>Multiquadrica</i>	1.4156 (cm)	5.0283 (cm)
	<i>Thin-Plate-Splines</i>	0.7907 (cm)	1.1306 (cm)

**Tabela 2** - Dados obtidos nos testes efectuados com os três modelos de imagem (Secção 2.1) para cada tipo de *RBF* (Secção 3.1).

RBF	Média	Desvio Padrão
<i>Gaussiana</i>	1.7377 (cm)	1.5887 (cm)
<i>Multiquadrica</i>	1.1487 (cm)	0.7614 (cm)
<i>Thin Plate Splines</i>	1.3397 (cm)	1.4262 (cm)

**Tabela 3** - Dados obtidos para o modelo de imagem catadioptrico com dois espelhos planos no período em que existem dois *viewpoints*. Os cálculos são referentes à distância entre o ponto estimado de intersecção das duas rectas e o ponto 3D dado pelo sistema de medição *Optotraker* (Secção 2.1) para cada tipo de *RBF* (Secção 3.1).



Modelo de Imagem	Pontos de Controlo (P)	Pontos de Calibração $\{y_i\}$	Pontos de Teste $\{w_i\}$
Projectivo	120	2880	824
Catadioptrico com Espelho Esférico	160	3120	652
Catadioptrico com dois Espelhos Planos	160	3120	1096

**Tabela 4** - Quantidade de pontos utilizados nas várias fases de experimentação.

## CAPÍTULO 7 - CONCLUSÕES/TRABALHO FUTURO

O método de calibração proposto por Miraldo *et al.* [1] pode ser utilizado com a generalidade das câmaras ou sistemas de câmaras, mesmo quando não existe um modelo analítico para esses sistemas. Ao necessitar de menos cálculos que as versões discretas do modelo genérico de imagem, a versão paramétrica tira também vantagem da redução de ruído devido à interpolação, contribuindo esta para tornar mais suave algum ponto mal marcado. Em vez de 7NM parâmetros para uma imagem com  $M \times N$  pixels, na versão com *raxel* simplificado, o método proposto precisa apenas de  $6P+18$  parâmetros.  $P$  é o número de pontos de controlo e é menor que  $M \times N$ . O método proposto necessita apenas de um ponto 3D para cada ponto na imagem, ao contrário das existentes onde eram necessários dois ou mais pontos 3D para cada ponto na imagem. É ainda possível identificar os diferentes centros de projecção no caso de sistemas centrais com várias câmaras ou vários espelhos. Mostrámos também que em sistemas com vários *viewpoints*, é possível estimar com bastante precisão um ponto 3D a partir de várias rectas. Como pode ser visto no Capítulo 6, foram obtidos erros de aproximadamente um centímetro, para uma distância média de três metros, em sistemas de múltiplos *viewpoints*, utilizando o método descrito no Secção 4.1.

Dos três sistemas analisados é fácil concluir que o projectivo foi o que obteve melhores resultados. Os resultados obtidos utilizando *thin-plate-splines* mostraram-se bastante melhores que os resultados com as *multiquadricas* e com as gaussianas. A ausência de parâmetro de forma e respectivo ajustamento revelou-se uma grande vantagem em relação aos outros dois tipos de *RBF* (Secção 3.1.1). Com erros de cerca de dois centímetros a uma distância média de três metros, para o caso do sistema catadioptrico com espelho esférico, consideramos que foram obtidos resultados muito positivos.

O modelo proposto por Miraldo *et al.* [1] apresenta resultados precisos. As *thin-plate-splines* melhoraram os resultados obtidos anteriormente e são muito mais simples de utilizar, por não precisarem de ajuste do parâmetro de forma. Desde que o número de pontos do dataset seja muito maior do que o número de pontos de controlo, não existem limitações para a calibração. Se os pontos do xadrez forem visíveis, o sistema não depende da resolução da câmara.

Como melhoramentos, seria interessante obter imagens mais nítidas e com maior resolução, no entanto, o sistema de sincronismo câmara/sistema de posicionamento *Optotraker*, limita o tipo de câmara a usar assim como a resolução das imagens adquiridas. Consideramos que existe bastante

espaço para testes com sistemas de câmaras mais complicados que os utilizados. Um sistema mais rápido, capaz de processar imagens com melhor qualidade, facilitaria a marcação e a precisão dos pontos na fase de tratamento da calibração.

## REFERÊNCIAS BIBLIOGRÁFICAS

1. P. Miraldo, H. Araujo e J. Queiró. *Point-based Calibration Using a Parametric Representation of the General Imaging Model*. In *International Conference on Computer Vision, 2011*. (Aceite para publicação)
2. M. Grossberg e S. Nayar. *A general imaging model and a method for finding its parameters*. In *International Conference on Computer Vision, 2001*.
3. M. Grossberg e S. Nayar. *The raxel imaging model and ray-based calibration*. *International Journal of Computer Vision*, 61:119–137, 2005.
4. S. Ramalingam, P. Sturm e S. Lodha. *Towards complete generic camera calibration*. In *Computer Vision and Pattern Recognition, 2005*.
5. P. Sturm e S. Ramalingam. *A generic concept for camera calibration*. In *European Conference on Computer Vision, 2004*.
6. R. Hartley e A. Zisserman. *Multiple view geometry in computer vision (2nd ed.)*. Cambridge University Press, The Edinburgh Building, Cambridge CB2 2RU, UK, 2000.
7. A. Agrawal, Y. Taguchi e S. Ramalingam. *Analytical forward projection for axial non-central dioptric & catadioptric cameras*. In *European Conference on Computer Vision*, Jan 2010.
8. J. Gluckman and S. Nayar. *Catadioptric Stereo Using Planar Mirrors*. *International Journal of Computer Vision* 44(1), 65–79, 2001
9. S. Ramalingam e P. Sturm. *Minimal solutions for generic imaging models*. In *Computer Vision and Pattern Recognition, 2008*.
10. H. Wendland. *Scattered Data Approximation*. Cambridge University Press, The Edinburgh Building, Cambridge CB28RU, UK, 2005.

11. T. Hastie, R. Tibshirani e J. Friedman. *The Elements of Statistical Learning*. Springer, 2001.
12. P. Miraldo, H. Araujo e J. Queiró. *Unique solution for the estimation of the plucker coordinates using radial basis functions*. In *International Conference on Computer Vision - Supplement Material*, 2011. (Aceite para publicação)
13. H. Pottmann e J. Wallner. *Computational Line Geometry*. Springer-Verlag, Berlin, 2001.
14. A. Bartoli and P. Sturm. *Structure-from-motion using lines: representation, triangulation, and bundle adjustment*. *Computer Vision and Image Understanding*, 100(3), 2005.
15. Northern Digital Incorporated. Optotrak certus motion capture system, 2009. <http://www.ndigital.com>