

**FEUP**

**MIEEC**  
**2010 – 2011**

**PDI**

Calibração automática de Sistema de Visão para aplicações  
em desporto

Aluno : Luís Manuel Queirós Martins – nº 040503120  
Coordenadores: Armando Jorge Sousa e Catarina Santiago

# Índice

1-Introdução	3
2-Estado da Arte	4
3-Objectivo do projecto	4
4-Âmbito do projecto	4
5-Modelo de Câmara PinHole	4
5.1-Parâmetros Intrínsecos	6
5.2-Parâmetros Extrínsecos	8
5.3-Tipo de Imagens Obtidas	8
6-Processamento de Imagem	8
6.1-Formação de imagem	9
6.2-Binarização de imagem	9
6.3-Determinação de pontos	10
6.3.1-Matching	10
6.3.2-Transformada de Hough	11
6.3.3-Algoritmo Vezhnevets	12
7-Processamento de Dados	13
8-Conclusões	14
9-Referencias	15

# 1-Introdução

A localização e/ou o mapeamento automático de pessoas e de objectos com o suporte de novas tecnologias são, na actualidade, importantes e, em diferentes circunstâncias, até essenciais. Vários sistemas, permitem ao utilizador controlar, gerir, detectar e prever situações que, normalmente, poderiam não ser observadas e dificilmente seriam objecto de monitorização. Esta tecnologia, face aos resultados positivos alcançados, é actualmente utilizada em múltiplas situações a exemplo das competições de Futebol Robótico, principalmente na denominada divisão dos robots mais pequenos (Small Size League) com a designação «5dpo». Normalmente no Futebol Robótico são utilizados sensores, câmaras e outro tipo de hardware normalmente implementados no corpo do robot, para permitir obter a localização em tempo real. Baseado nestas técnicas já utilizadas, um dos próximos passos poderá passar por projectar esta tecnologia em actividades da sociedade civil, nomeadamente em actividades desportivas.

Como é sabido, nestas actividades a coerência e a concentração das equipas de arbitragem são postas constante e sucessivamente em causa, permitindo com esta tecnologia, muito possivelmente, obter um contributo viável e válido para aumentar a objectividade das decisões, ajudando a actividade desportiva a manter-se credível assente em juízos imparciais. Numa fase posterior a este projecto, com estudo e investigação mais aprofundados neste âmbito, será possível fornecer a praticantes, treinadores, educadores, investigadores, decisores ou adeptos desportivos, entre os grupos mais relevantes, informações complementares e soluções suficientemente testadas. Será também relevante dizer que esta tecnologia poderá ainda ser utilizada em diferentes processos industriais ou no controlo do tráfego.

Neste projecto serão utilizadas câmaras portáteis, por serem não intrusivas aos sistemas, sendo as imagens obtidas utilizadas para localizar diversos pontos nos campos desportivos.

## 2-Estado da Arte

Vários métodos foram já estudados, no que diz respeito ao uso de sistemas de visão em desporto. Nomeadamente o uso de câmaras como sensores. Actualmente o sistema denominado Hawk-Eye provavelmente é o mais conhecido. Este sistema é muito conhecido pelo seu uso em vários torneios internacionais de ténis, nomeadamente no torneio do US-OPEN onde foi pela primeira vez implementado. O Hawk eye funciona com a utilização de multi-câmaras comparando a posição da bola pelas diferentes imagens obtidas. Este método foi estudado de tal modo que só era necessário utilizar imagens obtidas por *replay*. Este sistema desenvolvido por Sherry David e Hawkins Paul, investigadores britânicos, foi um resultado do estudo efectuado nesta área ao longo dos anos, conciliando o estudo de mecanismos de obtenção de imagens com os mecanismos de processamento de imagem. Actualmente está a ser ponderado o uso de cameras para o futebol de 11 com o objectivo de ajudar as equipas de arbitragem lances mais conflituosos.

## 3-Objectivo do projecto

Calculo automático dos parâmetros intrínsecos e extrínsecos de modelos de câmaras, utilizadas como sensores, do tipo pin hole através da detecção de pontos num campo de andebol ou num campo de futebol robótico graças ao processamento de imagem.

## 4-Âmbito do projecto

Este projecto terá como base de estudo principal as áreas científicas da visão artificial e um estudo matemático das projecções obtidas pelas câmaras.

## 5-Modelo de Câmara PinHole

Relativamente as câmaras pinhole que irão ser utilizadas, as imagens obtidas deverão ser semelhantes àquelas aqui reproduzidas:

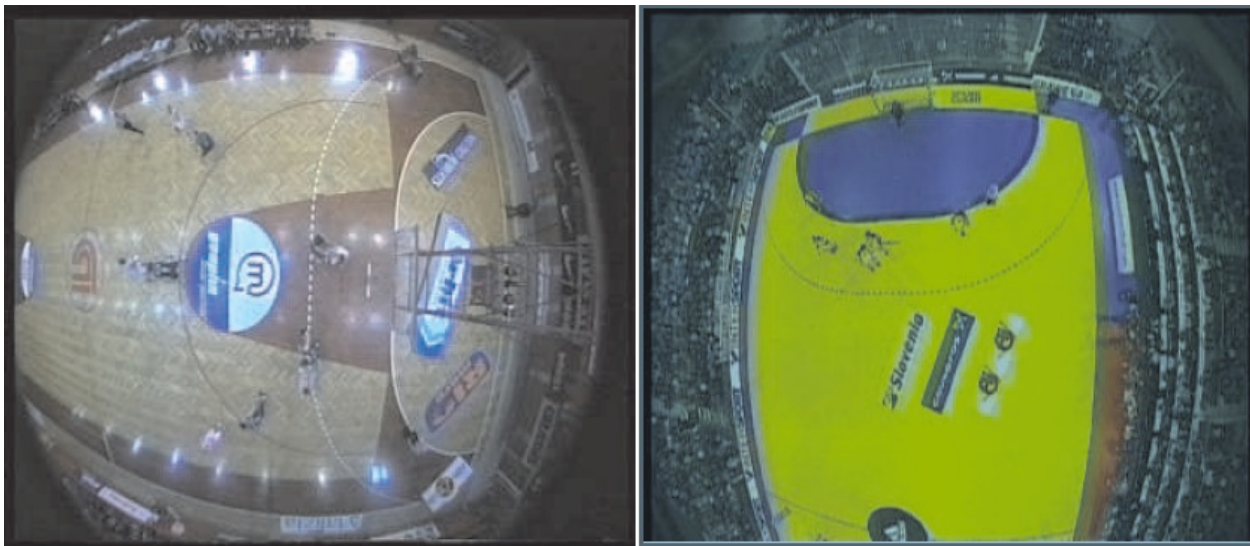


Figura 1 : Figura 2 – Fotos de campos de andebol por camera do tipo PinHole

Fonte: “Janez Pers, Matej Kristan, Matej Pers and Stanislav Kovaci - Analysis of Player Motion in Sport Matches -”

Com este tipo de câmaras sendo estas móveis, será necessário haver uma constante calibração dos parâmetros intrínsecos e extrínsecos que envolvem os sistemas observados. Os parâmetros referidos serão importantes para haver uma relação entre os pixeis das imagens e as coordenadas do campo observado. Caso não haja uma calibração correcta existira uma constante distorção radial nas imagens obtidas, principalmente observável se forem utilizadas 2 ou mais câmaras. Para se obterem imagens do campo inteiro observado conclui-se que será mesmo necessário a utilização de duas câmaras, ou seja, a calibração será definitivamente necessária para se obter resultados convincentes. Uma proposta deste projecto é obter as imagens em broadcast. A utilização de cameras em broadcast necessitará relacionar as imagens obtidas por cada câmara, sendo necessário então determinar até que ponto do campo cada câmara é capaz de observar. A vantagem do uso de câmaras em broadcast é a obtenção de imagens mais pormenorizadas das secções do campo.

As câmaras deverão se localizar, no caso do campo de andebol, entre a linha de nove metros (linha tracejado) e a linha de seis metros (linha continua), não sendo obrigatoriamente a posição da lente estar o 90º em relação ao solo. Na imagem seguinte está apresentado um exemplo da implementação das duas câmaras e as imagens:

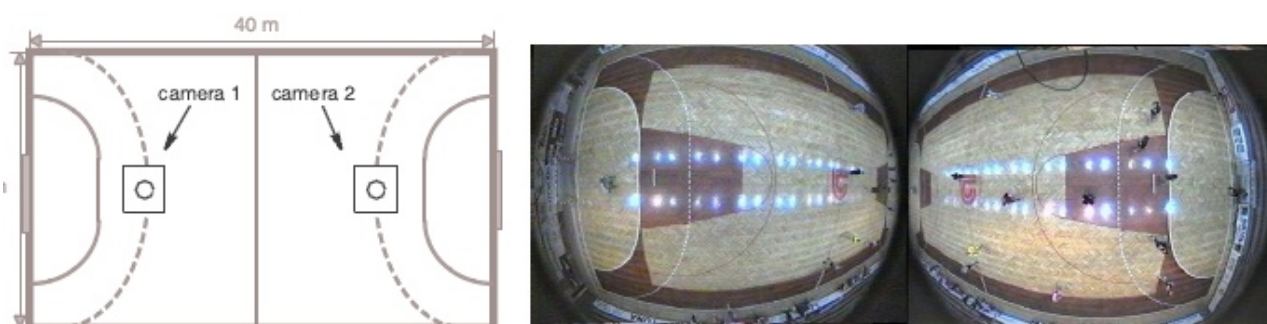


Figura 3 – Exemplos de posição espacial das cameras do tipo PinHole Model:

Figura 4 – Fotos de campos de andebol nas posições indicadas.

Fonte: “Janez Pers and Stanislav Kovaci - Tracking People in Sport: Making Use of Partially Controlled Environment”

O estudo matemático da calibração, demonstrado de seguida, deste tipo de câmaras, com o uso das imagens obtidas foi baseado por um trabalho do Pr. Dr. Armando Sousa [3].

### 5.1-Parâmetros Intrínsecos

A imagem obtida por qualquer câmara coloca a realidade tridimensional numa imagem bidimensional. Com a variância da localização, zoom e da amplitude da câmara é possível obter-se vários tipos de projecções.

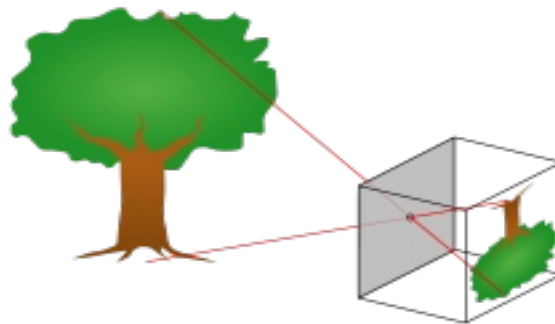


Figure 5: Esquema visual de uma câmara mostrando o plano virtual da imagem

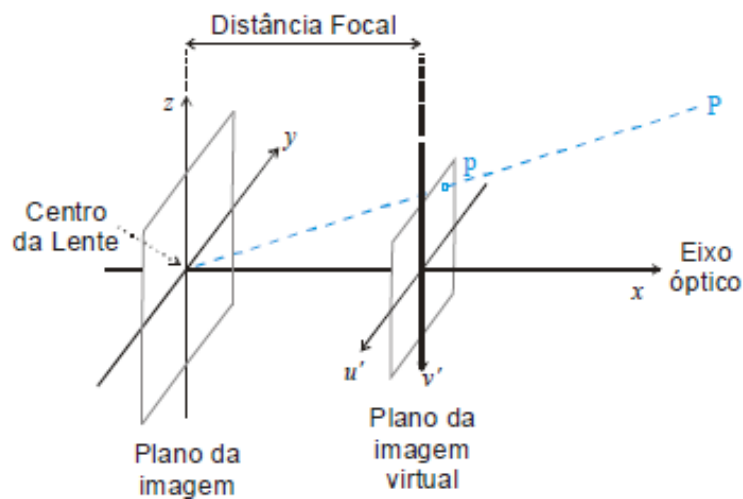


Figura 6 - Sistema de eixos no modelo da câmara

Fonte: "Armando Jorge Miranda de Sousa - Arquitecturas de Sistemas Robóticos e Localização em Tempo Real Através de Visão"

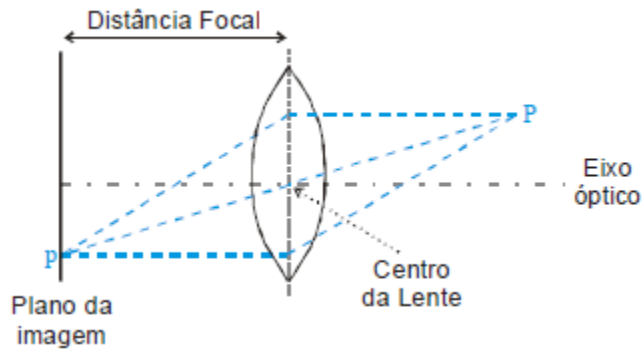


Figura 7 - Efeito de uma lente

Fonte: "Armando Jorge Miranda de Sousa - Arquiteturas de Sistemas Robóticos e Localização em Tempo Real Através de Visão"

A relação entre os pontos  $P(X,Y,Z)$  e os pontos da imagem  $(u',v')$  é:

$$\begin{cases} u' = \frac{-y}{x} * f \\ v' = \frac{-z}{x} * f \end{cases}$$

A distância focal da lente utilizada é a variável  $f$ . Utilizando um novo eixo de referencia para se considerar os pontos da imagem obtêm-se o sistema seguinte:

$$\begin{cases} u' = (u - O_u) * S_u \\ v' = (v - O_v) * S_v \end{cases}$$

No sistema anterior  $S_u$  é a distância horizontal entre pixels adjacentes e  $S_v$  a distância vertical entre pixels adjacentes.

Ao ser formada a imagem é provocado o efeito Barril que são distorções provocadas na imagem. Uma forma matemática para combater este efeito é o uso da seguinte expressão matemática:

$$\begin{cases} u' = (u_d - O_x) * S_u * (1 + k_b * r_d^2) \\ v' = (v_d - O_y) * S_v * (1 + k_b * r_d^2) \end{cases}$$

Na seguinte expressão  $u_d$  e  $v_d$  são as coordenadas da imagem distorcida e  $r_d$  é a distância ao centro da imagem. A variável  $k_b$  está relacionada com a distorção da lente.

Para ser calculado os parâmetros relativos à resolução da câmara irá ser utilizado a seguinte expressão:

$$\left\{ \begin{array}{l} Sx = axPx \\ Sy = Py \end{array} \right.$$

Sendo  $P_x$  e  $P_y$  as distancias dos pontos adjacentes e  $ax$  um factor de escala para compensar a eventualidade de haver frequências de amostragem iguais. Todos estes parâmetros após serem calculados não poderão ser alterados para o mesmo tipo de câmara e de óptica. Os parâmetros mencionados são:

- $O_u$  e  $O_v$
- $d$
- $k_b$
- $M_u$  e  $M_v$

## 5.2-Parâmetros Extrínsecos

Os parâmetros extrínsecos dependem unicamente da posição da câmara em relação ao exterior (mundo). Havendo uma enorme possibilidade de soluções as variáveis relativamente a estes parâmetros serão:

- $\beta$  Rotação em torno do eixo dos  $xx$ ;
- $\alpha$  Rotação em torno do eixo dos  $yy$ ;
- $\theta$  Rotação em torno do eixo dos  $zz$ ;
- $(X, Y, Z)$  Posição da câmara no mundo.

Estas variáveis serão calculadas no momento de implementação do sistema visual.

## 5.3-Tipo de Imagens Obtidas

As imagens obtidas poderão ser ou em fotografia ou em vídeo devendo estas não serem relativamente muito pesadas para não sobrecarregar o sistema de processamento desnecessariamente. Cada um dos resultados serão utilizados consoante o desenvolvimento do projecto. Sendo as fotografias testadas no inicio do projecto enquanto que os vídeos numa fase mais tardia devido à maior complexidade envolvente. Em princípio por facilidade de leitura e por serem ficheiros menos pesados para o sistema poderão ser do tipo .gif para as fotografias e do tipo .rmvb para os vídeos.

## 6-Processamento de Imagem

Nesta segunda parte do projecto o objectivo principal é a detecção de certos pontos nas imagens obtidas. Através de um algoritmo computacional, em linguagem C ou em C++ irão ser detectados pontos interessantes para o cálculo de parâmetros das câmaras utilizadas.

Os locais que deverão ser detectados para este projecto serão pontos intermedios, pontos de cruzamentos de linhas ( linha de lado com a linha atrasejado, linha do fundo com a linha de 6 metros) e o ponto de meio campo.



Para a detecção dos pontos será necessário recorrer a processamento de imagem no qual é necessário efectuar determinados passos para se obter o pretendido:

- *Formação de imagem*
- *Binarização de imagem*
- *Determinação dos pontos pretendidos*

### **6.1-Formação de imagem**

As imagens obtidas através das camaras serão a cores portanto para o processamento de imagem estas terão que ser passadas para preto e branco. Para tal será utilizado Matlab que tem comandos que permitem transformar as imagens analisadas variando os seus níveis de luz, cores, coordenadas espaciais entre outras propriedades. Estas técnicas serão essenciais para os passos seguintes porque os resultados desta fase permitirá facilitar na determinação dos pontos pretendidos. No caso da análise de imagens de campos de andebol será necessário obter imagens nas quais nestas deverão se destacar mais as linhas de fundo, laterais, 6 metros e a tracejado.

### **6.2-Binarização de imagem**

Para a binarização das imagens serão usadas técnicas de Thresholding. Depois de se obter uma imagem constituída por píxeis preto e branco deverá ser seleccionada posteriormente uma técnica que permitirá detectar regiões onde existirá uma razoável uniformidade nos valores das propriedades locais utilizadas, ou seja, haverá uma filtração do ruído da imagem binária.

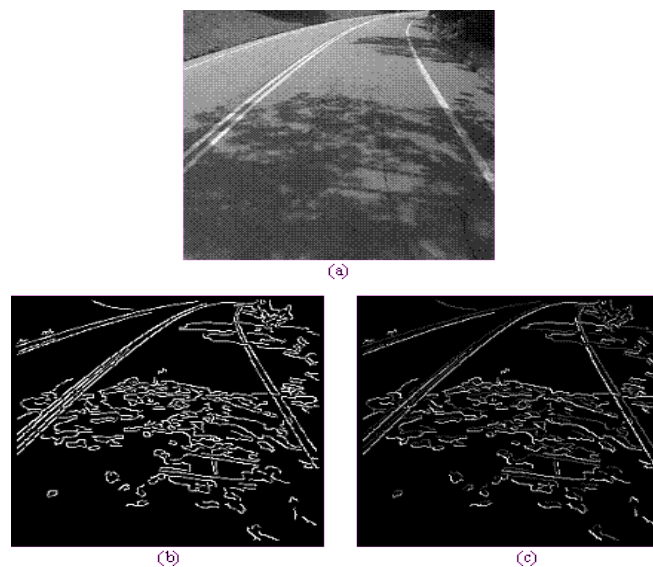


Figura 8 - Exemplo da binarização de uma imagem

Nas imagens anteriores (b e c) anterior pode-se observar os resultados obtidos depois da execução de dois tipos de thresholdings na imagem a). Conclui-se que com o uso de thresholding será possível obter-se uma filtragem de linhas facilitando assim a procura de locais pretendidos.

### 6.3-Determinação de pontos

Para a determinação dos pontos em campos de andebol será necessário a utilização de algoritmos para efectuar um varrimento da imagem para identificar os pontos. Existem várias técnicas de varrimento para a detecção de pontos ou linhas. Para este caso poderá ser utilizado, consoante os resultados pretendidos, os seguintes casos:

- *Matching para a detecção do meio campo*
- *Transformada de Hough para a detecção das linhas laterais e de fundo*
- *Algoritmo desenvolvido por Vezhnevets para o calculo da linha atrasejado*

#### 6.3.1-Matching

Esta técnica permite procurar nas imagens uma forma conhecida, estabelecendo uma relação entre o modelo predefinido e a imagem utilizada. O método consiste em deslocar o padrão para todas as posições possíveis e medir o grau de semelhança na área analisada. As medidas de dissemelhança mais comuns são:

$$\max_A |f-g| \quad \sum \sum_A |f-g| \quad \sum \sum_A (f-g)^2$$

Por expansão da potência, obter-se-á uma medida de semelhança:

$$\sum \sum_A fg = \frac{1}{2} [ \sum \sum_A f^2 + \sum \sum_A g^2 - \sum \sum_A (f-g)^2 ]$$

A expressão obtida é o coeficiente de correlação cruzada que com a desigualdade de Cauchy, verifica a relação seguinte:

$$C_{fg} = \sum \sum_A fg \leq \sqrt{(\sum \sum_A f^2 \cdot \sum \sum_A g^2)}$$

Quando a imagem e o padrão, na área de sobreposição, forem proporcionais a igualdade verifica-se. Ou seja para se encontrar a correspondência, é necessário encontrar as posições em que o coeficiente de correlação cruzada será máximo. Visto que o coeficiente de correlação depende da imagem deverá ser efectuada uma normalização que anule esta dependência.

$$\frac{C_{fg}}{\sqrt{\sum \sum_A g^2}}$$

Se a imagem e o padrão forem binarizados não será necessário o uso da normalização usando-se somente a expressão seguinte:

$$\Sigma \Sigma_A (f - f') \cdot g$$

A anterior expressão tem como função contar o número de 1's que coincidem com 1's do padrão proposto e subtrair o número de 1's coincidentes com os 0's do padrão. Os resultados obtidos são muito importantes porque é através destes que serão seleccionados os pontos da imagem que mais semelhanças têm com o modelo.

### 6.3.2-Transformada de Hough

As transformadas de Hough são métodos para detecção de linhas e curvas em que o espaço da imagem dos *pixels* é transformado num espaço de representação. Após a transformação dos *pixels* detectados a detecção das linhas faz-se no espaço transformado e por transformação inversa, obtém-se a expressão analítica das linhas do espaço-imagem original. Para a detecção de linhas considera-se um ponto P da imagem e a partir deste constrói-se um vector transformado. Ao longo deste vector quando se encontra um ponto, este irá ser incrementado num endereço dado por  $\omega = y_i \cdot x_i$ . Após a detecção de todos os pontos o vector construído apresentará contagens muito elevadas nos endereços correspondentes ao coeficiente angular das rectas passantes por P. Na imagem seguinte está um exemplo de como a transformada funciona:

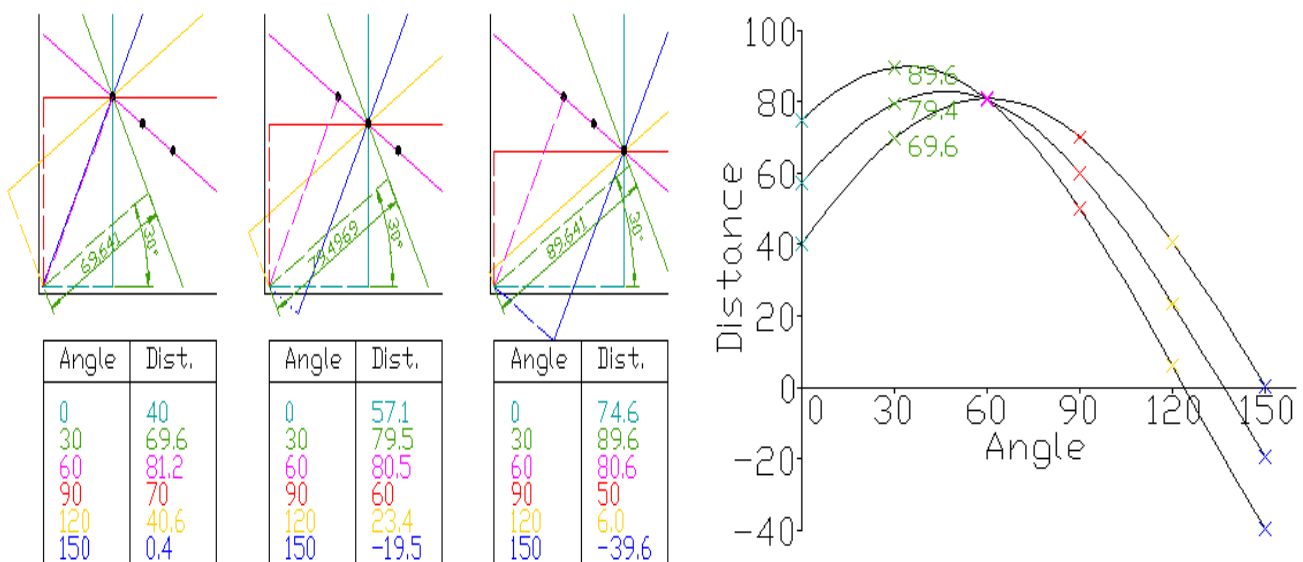


Figura 9 - Pontos Obtidos num vector transformado pelo método de Hough.  
 Figura 10 - Curvas Obtidas com os valores dos vectores obtidos

### 6.3.3-Algoritmo Vezhnevets

Este algoritmo foi desenvolvido com o intuito de detectar, em imagens com elevado ruído, conjuntos de quadrados de um tabuleiro de xadrez. Com uma configuração diferente dos parâmetros iniciais poder-se-á também utilizar este algoritmo para a detecção das linhas a tracejado dum campo de andebol visto que são formas aproximadamente quadradas.

O algoritmo inicialmente inclui um sistema próprio de binarização de imagem sendo utilizado um filtro thresholding adaptado que separa as zonas pretas das brancas. O inconveniente deste filtro é que as zonas quadradas que se pretendem localizar tendem a aumentar devido ao ruído existente ou à 'sujidade' da imagem. Devido a este inconveniente é utilizada uma aplicação de erosão originando assim uma separação dos cantos dos quadrados. Esta aplicação é bastante eficiente permitindo em imagens não uniformes obter resultados com um alto nível de segmentação. Nas figuras (Figura 10 e 11) a seguir verifica-se a actuação deste filtro:

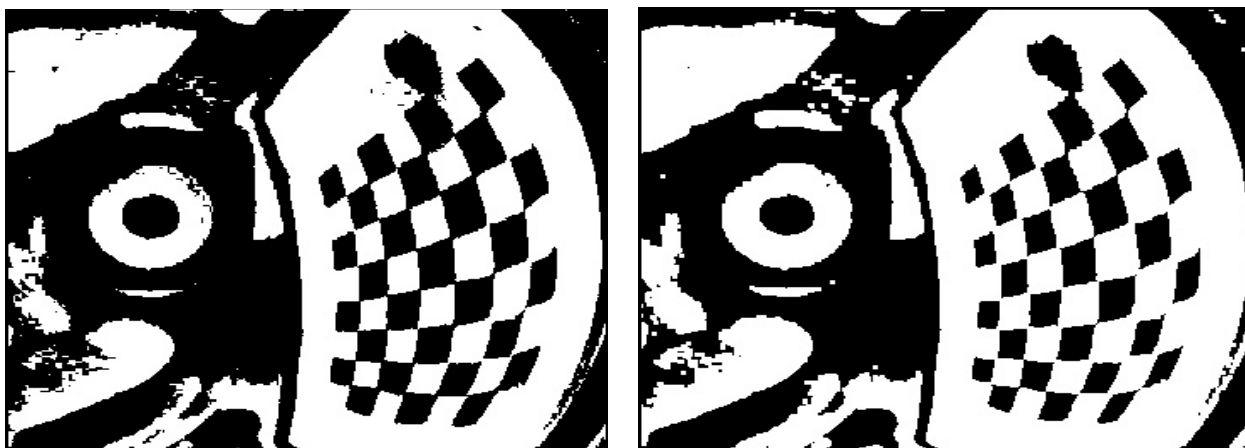


Figura 10 – Imagem com uma actuação na imagem original do filtro Thresholding Adaptado

Figura 11 – Imagem com duas actuações por parte do filtro Thresholding Adaptado

Fonte: “Martin Rufli, Davide Scaramuzza, and Roland Siegwart - Automatic Detection of Checkerboards on Blurred and Distorted Images ”

Após a obtenção duma imagem mais 'limpa' o próximo passo será identificar as zonas a tracejado. Será utilizado um gerador de quadrados que funciona como um contador binário em que localiza contornos próximos e após a localização seja positiva analisa esses píxeis e tenta-os ajustar aproximadamente a um polígono. No caso de não haver a localização total dos quadrados da imagem aplica-se uma nova erosão da imagem e volta-se a tentar localizar mais quadrados. Após estar concluída a procura de todos ou quase todos os quadrados da imagem aplica-se um próximo passo. A função do passo seguinte é juntar os quadrados obtidos através de links. Este passo não será necessário neste projecto visto que só se pretende é localizar a linha tracejada. Na figura 12 a seguir mostrada observa-se a localização de alguns quadrados com a actuação de diferentes passos de erosão:

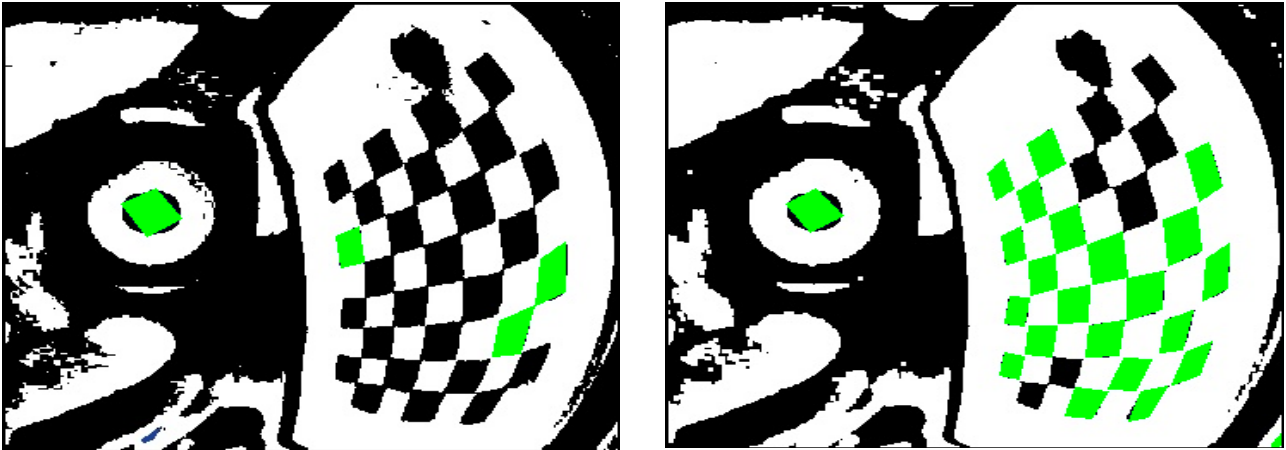


Figura 10 – Primeira Erosão e quadrados encontrados

Figura 11 – Segunda Erosão e quadrados encontrados

Fonte: “Martin Rufli, Davide Scaramuzza, and Roland Siegwart - Automatic Detection of Checkerboards on Blurred and Distorted Images”

## 7-Processamento de Dados

Neste projecto a área relacionada com o processamento de dados está relacionada intimamente com o processamento de imagens visto que obteremos através dos resultados das imagens os dados necessários para que através dos parâmetros intrínsecos da câmara se possam determinar os pontos do campo que esta consegue observar. Em caso de serem utilizadas mais do que uma câmara para o visionamento vai ser efectuada uma leitura dos parâmetros de todas as câmaras envolvidas e calibra-las de tal modo que seja possível observar o campo inteiro. Os resultados obtidos deverão ser os seguintes indicados nas figuras 12 e 13:

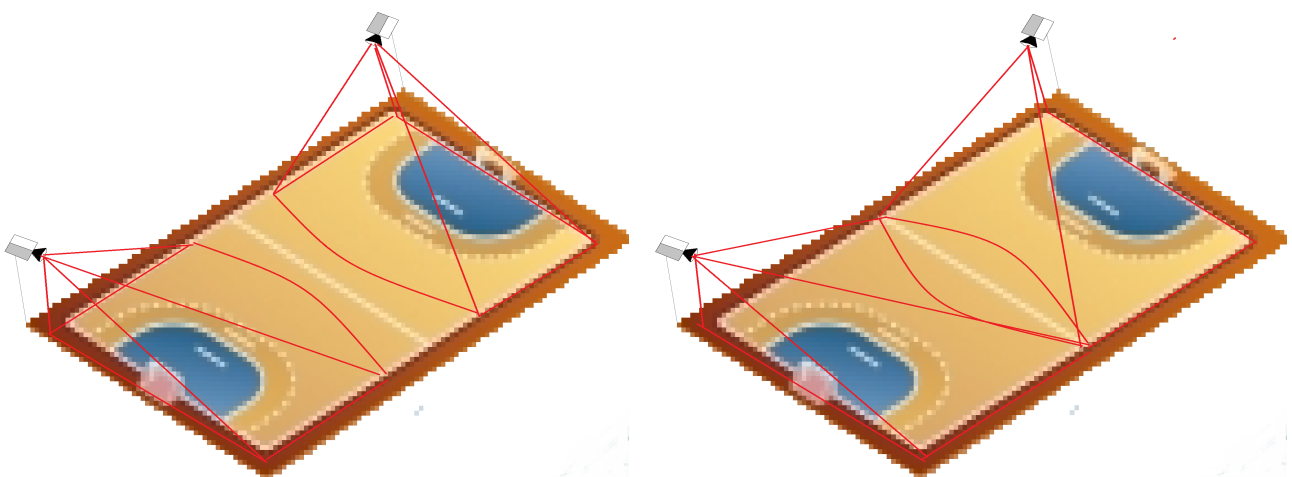


Figura 12– Camaras sem calibração

Figura 13 – Camaras com calibração

## 8-Conclusões

Este trabalho será então desenvolvido principalmente à volta das áreas de processamento de imagem e da calibração dos mecanismos de obtenção de imagem. Todas estas áreas são muito promissoras no futuro, tanto para ambientes fabris, sociais, desportivos entre outros sendo uma necessidade futura.

Com o desenvolvimento nestas áreas será possível obterem-se automaticamente informações não observadas à primeira vista ou então de maneiras mais eficazes e mais rápidas O processamento de imagem é então uma área promissora com uma enorme flexibilidade para qualquer área

A calibração de câmaras por meios informáticos além de permitir ao utilizador obter melhores imagens permitirá o uso de câmaras móveis em qualquer local predefinido.

## 9-Referencias

- [1] – Artigo “ Janez Pers and Stanislav Kovaci - Tracking People in Sport: Making Use of Partially Controlled Environment ”
- [2] – Artigo “ Janez Pers, Matej Kristan, Matej Pers and Stanislav Kovaci - Analysis of Player Motion in Sport Matches ”
- [3]-Artigo “Martin Ruffli, Davide Scaramuzza, and Roland Siegwart - Automatic Detection of Checkerboards on Blurred and Distorted Images ”
- [4] – Tese de Doutoramento “ Armando Jorge Miranda de Sousa - Arquiteturas de Sistemas Robóticos e Localização em Tempo Real Através de Visão ”
- [5] – Artigo “Catarina Santiago, Armando Sousa, Maria Luísa Estriga, Luis Paulo Reis e Martin Lames - Survey on Team Tracking Techniques Applied to Sports”
- [6] – “Y. R. Tsai. A versatile camera calibration technique for high-accuracy 3d machine vision metrology using off-the-shelf tv cameras and lenses. IEEE Journal of Robotics and Automation, RA-3(4):323–344, 1987.”
- [7] – Artigo “Hongzhi Wang e John Oliensis – Shape Matching by segmentation Averaging”
- [8] – Livro “Gonzalez, R.C. and Woods, R.E. – Digital Image Processing, Prentice Hall, 1993”
- [9] – Artigo “Li.S.Y, Ruan.X.F,Zhu.Q.H - The research of vision calibration method for target localization”
- [10] – Artigo “ J.Apolinar Muñoz Rodríguez - Laser imaging and approximation networks for calibration of three-dimensional vision “
- [11] – Artigo “Michel Dome, Marc Richetin, Jean-Thierry Lapreste and Gerard Rives -Determination of the Attitude of 3-D Objects from a Single Perspective View”
- [12] – Conteúdos Académicos da disciplina de Sistemas Baseados em Visão - FEUP