

Sistema de Localização Visual da Equipe de Futebol de Robôs POTI-UFRN (versão 2008) na Categoria Very Small Size

Ellon P. MENDES*, Adelardo A.D. MEDEIROS*

*UFRN-CT-DCA

59072-900 Natal RN

Email: [ellonpaiva,adelardo]@dca.ufrn.br

Abstract—Este artigo descreve o sistema de localização empregado na versão 2008 da equipe POTI de futebol de robôs, categoria Very Small Size. São apresentadas as técnicas de calibração de câmera e o sistema de cor utilizado (HPG) para segmentação das regiões de cor na imagem. Mostram-se resultados experimentais que comprovam a eficiência da abordagem proposta.

I. INTRODUÇÃO

A equipe POTI foi desenvolvida no DCA-UFRN para a categoria Very Small Size de futebol de robôs. Uma equipe é composta por três robôs (figura 1) e outros sub-sistemas de hardware e software, conforme figura 2. No nosso caso, cada robô é identificado por um rótulo com a cor da equipe (azul ou amarelo) e um rótulo auxiliar para individualização.

O sistema de visão consiste em uma câmera que capta imagens do campo com os robôs e a bola. O sistema de localização fornece a posição dos robôs e da bola no campo através de marcas coloridas colocadas nos robôs. O sistema de estratégia gera a posição futura dos robôs a partir das posições atuais adquiridas e do comportamento desejado. O sistema de controle gera o sinal que vai efetivamente deslocar o robô, de forma que ele chegue à posição definida pela estratégia. Finalmente, o sistema de transmissão envia os sinais de controle para o robô.



Fig. 1. Robôs da equipe POTI-UFRN

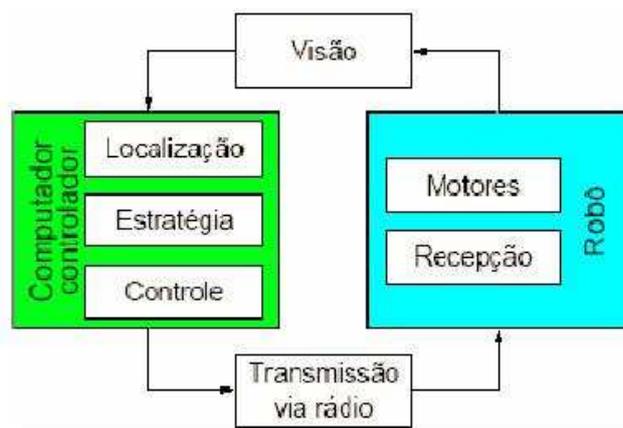


Fig. 2. Arquitetura do sistema de futebol de robôs

Neste artigo será focalizado o sistema de localização, que foi o principal aprimoramento realizado na equipe POTI no período de 2007-2008.

II. LOCALIZAÇÃO

O sistema de localização tem como objetivo extrair a posição dos robôs e da bola em campo a partir de uma imagem adquirida de uma câmera. No nosso sistema, foram utilizadas câmeras Firewire 400 (IEEE 1394a) da TheImagingSource [1], modelos DBK 21AF04 (640×480, 60fps) e DFK 41BF02 (1280×960, 15fps).

Quando uma nova imagem é recebida, o sistema de localização faz uma varredura procurando píxeis azuis, amarelos ou laranjas, que são as cores obrigatórias das duas equipes e da bola. Como os objetos têm tamanho mínimo conhecido, a busca é feita apenas em alguns píxeis, a cada N linhas e colunas (no caso, $N = 5$).

Quando um píxel de uma das cores principais é encontrado, buscam-se todos os píxeis vizinhos da mesma cor, usando uma versão otimizada do algoritmo *seedfill* [2]. É feita então uma média das coordenadas dos píxeis vizinhos de mesma cor, o que define o centro da região. Caso se trate de uma região azul ou amarela, é feita uma varredura nas proximidades para localizar regiões de cores correspondentes aos rótulos auxiliares. O centro do robô e sua orientação são calculados a partir dos centros das regiões que correspondem aos seus dois rótulos.

Uma vez calculados o centro da região que corresponde à bola e os centros dos robôs, estas informações expressas em unidades de píxeis são convertidas para metros, utilizando dados de calibração da câmera.

As linhas gerais desta forma de localizar os objetos no campo são conhecidas e utilizadas pela maioria das equipes que disputam esta categoria de competição. Os diferenciais da nossa abordagem se concentram em dois aspectos:

- Etiquetagem dos píxeis, ou seja, a forma de associar um píxel descrito pelas suas componentes RGB com uma determinada cor.
- Conversão das coordenadas na imagem (em píxeis) para coordenadas no campo de jogo (em metros).

III. ETIQUETAGEM DE PÍXEIS

O sistema de localização se utiliza de um modelo de cores desenvolvido especificamente para esta aplicação, denominado HPG. O HPG é apropriado para aplicações onde seja necessário distinguir entre regiões de cor e regiões de cinza, como na categoria Very Small Size de futebol de robôs.

A. Modelo de Cores HPG

O nosso sistema utiliza o modelo de cores denominado *Hue-Purity-Grayness*, ou simplesmente HPG. Este modelo foi desenvolvido no DCA-UFRN para ser aplicado no futebol de robôs e em outras aplicações onde é necessário reconhecer regiões de cores específicas e diferenciá-las de regiões em tons de cinza. No HPG, um píxel é definido como sendo composto por uma componente de cor pura e por uma componente de tom de cinza puro, ponderados por um fator de pureza P .

Sejam $[R \ G \ B]^T$ as componentes vermelho, verde e azul de um píxel e sejam X , D e N definidos por:

$$\begin{aligned} X &= \max(R, G, B) / 255 \\ D &= \text{med}(R, G, B) / 255 \\ N &= \min(R, G, B) / 255 \end{aligned} \quad (1)$$

A equação que define os valores de P e G para um píxel é dada por:

$$\begin{bmatrix} X \\ D \\ N \end{bmatrix} = P \cdot \begin{bmatrix} 1 \\ c \\ 0 \end{bmatrix} + (1 - P) \cdot \begin{bmatrix} G \\ G \\ G \end{bmatrix} \quad (2)$$

Da equação 2 temos que:

$$P = X - N \quad G = \frac{N}{1 - (X - N)} \quad c = \frac{(D - N)}{(X - N)}$$

O valor de H pode ser obtido por:

$$H = \begin{cases} \text{INDEFINIDO} & \text{se } X = N \\ \frac{\pi}{3} \cdot (f + c) & \text{se } f = 0, 2, -2 \\ \frac{\pi}{3} \cdot (f + (1 - c)) & \text{se } f = 1, -1, -3 \end{cases} \quad (3)$$

onde f indica a faixa onde se encontra o píxel no círculo de cores, mostrado na figura 3

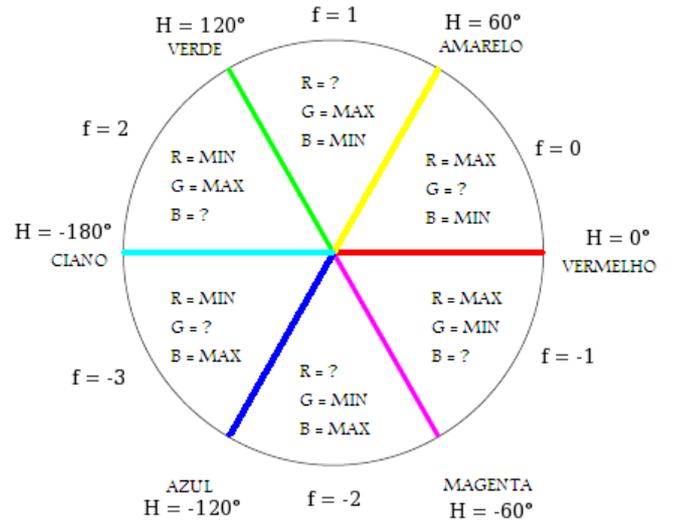


Fig. 3. Círculo de Cores

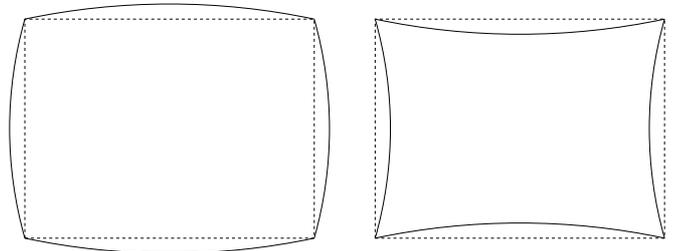


Fig. 4. Tipos de distorção radial: barril e ampulheta

IV. CÁLCULO DA POSIÇÃO

No sistema de localização, precisamos obter a posição do robô em metros a partir da sua posição na imagem. Para tal, primeiramente retiramos o efeito da distorção radial na imagem causado pela lente e depois convertemos efetivamente as coordenadas do ponto em metros utilizando uma matriz de homografia.

A. Distorção Radial

A distorção radial é um efeito causado na imagem devido a imperfeições no sistema de lentes da câmera. A figura 4 mostra as distorções dos tipos barril e ampulheta.

Sejam:

$$Q = \begin{bmatrix} u^D \\ v^D \end{bmatrix} \quad P = \begin{bmatrix} u \\ v \end{bmatrix} \quad C = \begin{bmatrix} U^C \\ V^C \end{bmatrix}$$

onde u^D e v^D são as coordenadas de um ponto distorcido, u e v as coordenadas do ponto correspondente sem distorção e U^C e V^C as coordenadas do centro de distorção radial. Pode-se modelar o efeito de distorção radial [3] por:

$$P = (1 + kr^2) \cdot Q \quad (4)$$

onde k é o coeficiente de distorção radial e r é a distância (raio) do ponto distorcido em questão ao centro de distorção radial C :

$$r^2 = \|Q - C\|^2 = (u^D - U^C)^2 + (v^D - V^C)^2 \quad (5)$$

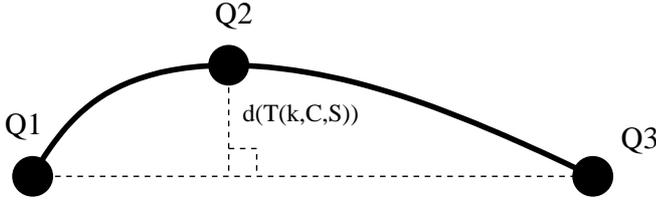


Fig. 5. Cálculo de $d(T(k, C, S))$

Excluindo-se o efeito causado pela distorção radial, todas as transformações de câmera realizadas para transformar pontos em coordenadas de mundo para coordenadas de imagem preservam a propriedade da colinearidade. Ou seja, caso existam três ou mais pontos que pertençam a uma reta no mundo, os pontos correspondentes estarão contidos em uma reta na imagem, se não houver distorção radial. Logo, é possível quantizar a distorção radial através da medida do desalinhamento de três ou mais pontos na imagem correspondentes a pontos de mundo que se sabe pertencerem a uma reta.

Se:

$$S = \langle Q^1, Q^2, Q^3 \rangle \quad (6)$$

é um conjunto de pontos na imagem distorcida, correspondentes a pontos pertencentes a uma reta no mundo, a função:

$$\begin{aligned} T(k, C, S) &= \langle P^1, P^2, P^3 \rangle \\ &= \langle (1+k\|Q^1-C\|) \cdot Q^1, \\ &\quad (1+k\|Q^2-C\|) \cdot Q^2, \\ &\quad (1+k\|Q^3-C\|) \cdot Q^3 \rangle \end{aligned} \quad (7)$$

calcula o conjunto de pontos $\langle P^1, P^2, P^3 \rangle$ corrigidos, a partir de S , k e C . Podemos mensurar a distorção $d(T(k, C, S))$ como sendo a distância entre o ponto intermediário e a reta formada pelo pontos pontos extremos. A figura 5 mostra um exemplo do cálculo de $d(T(k, C, S))$ com Q^1 e Q^3 como pontos extremos e Q^2 como ponto intermediário. Este cálculo nos mostra o quanto de distorção restou após a correção utilizando um certo k e um certo C . Note que no caso dos pontos estarem alinhados, o valor retornado é zero. Considerando o conjunto de tuplas S_1, \dots, S_N , pode-se calcular o erro total causado pela distorção radial como:

$$E(k, C) = \sum_{i=1}^N d(T(k, C, S_i)) \quad (8)$$

e finalmente, podemos calcular:

$$\left[\hat{k}, \hat{C} \right] = \min_{k, C} E(k, C) \quad (9)$$

onde \hat{k} e \hat{C} são, respectivamente, o coeficiente de distorção radial e o centro de distorção radial que melhor corrigem a imagem distorcida.

Para realizar a correção no sistema de localização, são gerados vários conjuntos de três pontos na imagem capturada, cada conjunto pertencente a uma reta conhecida.

Estas retas podem ser obtidas facilmente localizando na imagem os pontos notáveis no campo, como por exemplo as bordas ou a linha central. Então, pode-se utilizar algum dos algoritmos de otimização conhecidos para encontrar os valores de \hat{k} e \hat{C} que minimizam $E(k, C)$. Uma vez calculados \hat{k} e \hat{C} , pode-se calcular qualquer ponto corrigido P a partir de um ponto distorcido Q utilizando a equação 4. No nosso sistema, a minimização foi realizada utilizando algoritmos genéticos.

B. Homografia

Para que se possa usar uma câmera como sensor de posição no futebol de robôs, é preciso que se possa extrair, a partir das coordenadas de um píxel na imagem, as coordenadas do ponto correspondente no sistema de coordenadas do mundo. Desta maneira, pode-se então identificar os robôs e a bola através dos algoritmos de detecção de regiões e determinar as suas posições correspondentes no campo.

No futebol de robôs em particular, a posição de mundo que nos interessa é planar. Sabe-se [4] que a relação entre as coordenadas de um ponto em um plano no mundo e as coordenadas deste mesmo ponto em uma imagem são relacionadas por uma matriz de homografia, através da relação:

$$s \cdot \bar{\mathbf{B}}^I = \mathbf{H} \cdot \bar{\mathbf{Q}}^W \quad (10)$$

onde:

$$\bar{\mathbf{B}}^I = \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \quad \bar{\mathbf{Q}}^W = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

\mathbf{H} é uma matriz (3×3) , conhecida como *homografia*, e s um fator de normalização.

Denominando-se \mathbf{h}_1 , \mathbf{h}_2 e \mathbf{h}_3 as três linhas da matriz \mathbf{H} , a equação 10 pode ser reescrita:

$$s \cdot \bar{\mathbf{B}}^I = \begin{bmatrix} \mathbf{h}_1 \\ \mathbf{h}_2 \\ \mathbf{h}_3 \end{bmatrix} \cdot \bar{\mathbf{Q}}^W \quad (11)$$

e logo:

$$s = \mathbf{h}_3 \cdot \bar{\mathbf{Q}}^W \quad (12)$$

substituindo 12 em 11 temos:

$$(\mathbf{h}_3 \cdot \bar{\mathbf{Q}}^W) \bar{\mathbf{B}}^I = \begin{bmatrix} \mathbf{h}_1 \\ \mathbf{h}_2 \\ \mathbf{h}_3 \end{bmatrix} \cdot \bar{\mathbf{Q}}^W \quad (13)$$

A matriz \mathbf{H} multiplicada por um escalar também será uma solução da equação 11. Para que a homografia tenha um valor único, fixaremos $h_{33} = 1$.

Tomando-se as equações que representam as duas primeiras linhas da equação vetorial 11, já que a terceira linha é sempre verdade por construção, chega-se a:

$$\begin{cases} u^I x^W \cdot h_{31} + u^I y^W \cdot h_{32} + u^I = \\ \quad x^W \cdot h_{11} + y^W \cdot h_{12} + h_{13} \\ v^I x^W \cdot h_{31} + v^I y^W \cdot h_{32} + v^I = \\ \quad x^W \cdot h_{21} + y^W \cdot h_{22} + h_{23} \end{cases} \quad (14)$$

As duas equações do sistema 14 foram obtidas a partir de um único ponto na imagem. Havendo n pontos, o conjunto de $2n$ equações pode ser representado matricialmente por:

$$\mathbf{Ax} = \mathbf{b} \quad (15)$$

onde:

$$A = \begin{bmatrix} x_1^W & y_1^W & 1 & 0 & 0 & 0 & -u_1^I x_1^W & -u_1^I y_1^W \\ 0 & 0 & 0 & x_1^W & y_1^W & 1 & -v_1^I x_1^W & -v_1^I y_1^W \\ x_2^W & y_2^W & 1 & 0 & 0 & 0 & -u_2^I x_2^W & -u_2^I y_2^W \\ 0 & 0 & 0 & x_2^W & y_2^W & 1 & -v_2^I x_2^W & -v_2^I y_2^W \\ \vdots & \vdots \\ x_n^W & y_n^W & 1 & 0 & 0 & 0 & -u_n^I x_n^W & -u_n^I y_n^W \\ 0 & 0 & 0 & x_n^W & y_n^W & 1 & -v_n^I x_n^W & -v_n^I y_n^W \end{bmatrix}$$

$$x = \begin{bmatrix} h_{11} \\ h_{12} \\ h_{13} \\ h_{21} \\ h_{22} \\ h_{23} \\ h_{31} \\ h_{32} \end{bmatrix} \quad b = \begin{bmatrix} u_1^I \\ v_1^I \\ u_2^I \\ v_2^I \\ \vdots \\ u_n^I \\ v_n^I \end{bmatrix}$$

O sistema 15 tem $2n$ equações e 8 incógnitas. Se há 3 ou menos pontos na imagem para os quais se conhece a posição no mundo, é impossível calcular a homografia. Com exatamente 4 pontos, o vetor \mathbf{x} com os elementos da matriz \mathbf{H} é determinado por:

$$\mathbf{x} = \mathbf{A}^{-1} \mathbf{b} \quad (16)$$

Com $n > 4$, o sistema deve ser resolvido de forma a minimizar o erro quadrático médio, através da pseudo-inversa:

$$\mathbf{x} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{b} \quad (17)$$

Para calcular a homografia, são obtidos pontos na imagem referentes a pontos notáveis no campo (quinas e pontos de cruzamento de linhas) durante a fase de calibração. Os pontos de mundo e seus correspondentes na imagem são utilizados no sistema 15. Como existe ruído adicionado na determinação dos pontos notáveis na imagem, a qualidade do resultado obtido será tanto melhor quanto maior for o número de pontos utilizados.

V. RESULTADOS

O modelo HPG mostrou-se bastante eficiente na detecção das marcas coloridas no campo. Com apenas um ajuste de P, foi possível distinguir entre píxeis coloridos e píxeis pertencentes ao campo. A figura 6 mostra:

- A imagem utilizada para os testes
- A imagem com os píxeis abaixo de um limiar de P (tons de cinza)
- A imagem com os píxeis acima de um limiar de P (cores). Nesta imagem, podem ser aplicados limiares de H para identificar os objetos.
- O resultado da aplicação de um limiar de G na imagem de tons de cinza.

Para o cálculo da distorção radial foram utilizados vários conjuntos de três pontos pertencentes a retas conhecidas no campo. Os pontos utilizados são mostrados na figura 7, onde também se mostra o resultado da correção da distorção. Na imagem corrigida, percebe-se que o plano do campo e o plano da imagem não estão perfeitamente paralelos, o que indica que há um certo desalinhamento da câmera. Este fato não traz maiores prejuízos para o desempenho do sistema, tendo em vista que o cálculo da posição feito através da homografia já leva em conta este efeito.

REFERENCES

- [1] www.theimagingsource.com
- [2] R. C. Gonzalez, R. E. Woods. *Digital Image Processing*. Prentice Hall, 2ª Edição, 2002.
- [3] E. Trucco, A. Verri. *Introductory Techniques for 3-D Computer Vision*. Prentice Hall, 1998.
- [4] D. A. Forsyth, J. Ponce. *Computer Vision: A Modern Approach*. Prentice Hall, 2002.

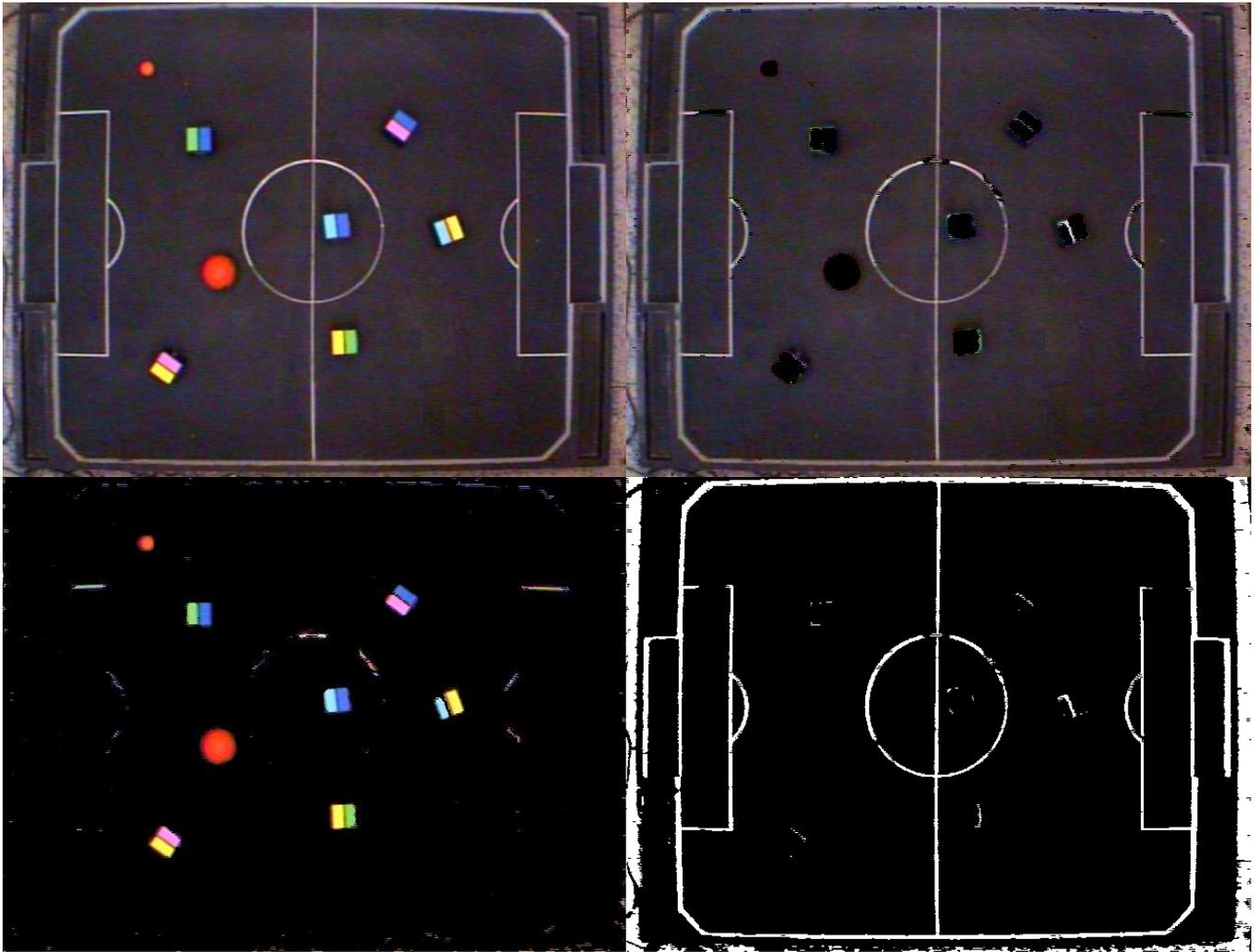


Fig. 6. Processamento usando modelo HPG: imagem original, só os píxeis sem cor, só os píxeis coloridos e os píxeis das linhas do campo

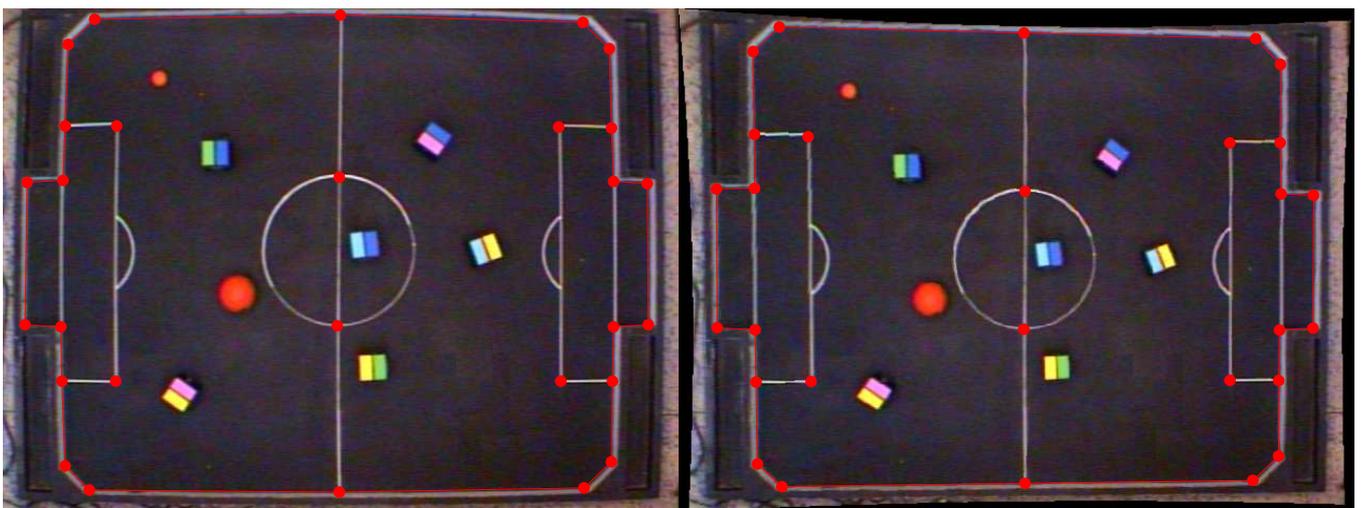


Fig. 7. Imagens com e sem distorção radial