

## Calibração de múltiplas câmeras baseado em um padrão invariante

Manuel E. Loaiza, Alberto B. Raposo, Marcelo Gattass

Tecgraf, Dept. of Informatics  
PUC-Rio - Brazil

{manuel, abraoso, mgattass}@tecgraf.puc-rio.br

### Resumo

*Neste artigo apresentamos um método de calibração de múltiplas câmeras baseado na extração e posterior uso das características físicas e implícitas de um padrão unidimensional invariante, cuja estrutura é definida por 4 marcadores colineares. As vantagens deste padrão de calibração aparecem em duas etapas-chave do processo de calibração: na captura eficiente de uma amostra robusta de pontos de referência para a calibração inicial, e na otimização dos parâmetros de calibração das câmeras. Apresentamos também os resultados obtidos calibrando dois sistemas de múltiplas câmeras usando o método proposto e comparando-os com outros métodos propostos na literatura.*

### Abstract

*In this paper we present a method for multiple cameras calibration based on the extraction and use of physical characteristics of an unidimensional invariant pattern, which structure is defined by 4 colinear markers. The advantages of this kind of pattern appears in two key steps of the calibration process: the efficient capture of a robust sample of reference points for the initial calibration, and the optimization of camera calibration parameters. We also present the results obtained by the calibration of two multiple cameras systems using the proposed method and compare them with other methods proposed in the literature.*

### 1. Introdução

Na área de realidade virtual e aumentada, um componente chave tem sido o sistema de rastreamento, o qual permite criar uma interface de interação entre o usuário e o cenário virtual. Para implementar esse tipo de sistema diversas tecnologias foram usadas como: sonora, eletromagnética, mecânica, óptica e híbrida. Porém, é a tecnologia óptica a mais utilizada principalmente por

duas características principais: disponibilidade e baixo custo dos equipamentos necessários para montar um sistema desse tipo, e a característica de não poluição do campo de rastreamento com objetos como cabos e outros componentes mecânicos e eletrônicos.

Nos últimos anos, um dos principais objetivos na implementação dos sistemas de rastreamento óptico é permitir o aumento do campo de rastreamento sem perder precisão na captura e a recuperação da posição e da orientação 3D dos objetos rastreados. Para cumprir tal objetivo, em termos de equipamento, estes sistemas têm incrementado o número de câmeras que os compõem, com a intenção de capturar com boa nitidez um maior espaço real de rastreamento.

Já na parte teórica, as pesquisas caminham na direção de aprimorar cada vez mais o processo de calibração das câmeras levando em conta o gerenciamento simultâneo das múltiplas câmeras. Foi assim que o processo de calibração das câmeras ganhou importância e é considerado o coração do funcionamento dos sistemas de rastreamento óptico. A precisão alcançada no cálculo da posição 3D dos objetos rastreados está diretamente relacionada a uma correta e precisa calibração das câmeras em relação à área comum de rastreamento. Seguindo essa direção, os diferentes métodos de calibração vêm explorando novas formas de aprimorar o processo de calibração, uma delas aponta para a pesquisa de novos padrões sintéticos para calibração, onde se requer que eles possuam algumas características básicas como:

- Ter um formato simples que permita seu fácil reconhecimento e extração a partir das imagens capturadas pelas câmeras.
- Ser nitidamente visíveis em diferentes perspectivas em relação a posição de duas a mais câmeras.
- Permitir uma fácil manipulação e liberdade de movimento.
- Oferecer suficiente informação implícita a seu formato que permita levar adiante um preciso processo de calibração das câmeras.

É pensando nessas características que o método proposto neste artigo apresenta como primeira contribuição o uso de um padrão 1D invariante à projeção. Este padrão permite a captura e recuperação de uma boa quantidade de pontos de referência de maneira robusta, eficiente e simultaneamente nas diferentes câmeras do sistema de rastreamento. A segunda contribuição está na exploração das características físicas do padrão invariante, que ajudam a aprimorar o cálculo dos parâmetros da calibração.

Na seção 2, faremos uma descrição básica das etapas que compõem o processo de calibração de múltiplas câmeras, assim como os padrões de calibração comumente usados. Na seção 3, descreveremos o método proposto neste artigo. Na seção 4 apresentaremos nosso ambiente de teste e resultados. Na seção 5 apresentaremos nossas conclusões e proposta de trabalhos futuros.

## 2. Calibração de múltiplas câmeras

Nesta seção descrevemos alguns conceitos relacionados ao processo de calibração de múltiplas câmeras, e para guiar essa descrição propomos um **framework conceitual** que permitirá uma melhor organização das etapas comumente encontradas nos diferentes métodos de calibração de múltiplas câmeras.

A calibração de múltiplas câmeras é o processo pelo qual os parâmetros intrínsecos e extrínsecos de um conjunto de duas ou mais câmeras são calculados simultaneamente. Para isso, precisa-se de uma amostra de pontos visíveis e corretamente correspondidos no espaço de interseção dos diferentes campos de visão definidos pela posição e orientação das câmeras.

O framework conceitual definido aqui engloba todos os processos de calibração, desde a aquisição dos dados para calibração das câmeras até o refinamento e otimização dos parâmetros calculados. A proposta da estrutura de nosso framework foi definida pela análise de diferentes métodos analisados como os descritos em: [2][3][14][17][13][9]. O framework é definido por três etapas, que serão descritas nas subseções seguintes.

### 2.1 Primeira etapa: Aquisição de dados

A primeira etapa do processo de calibração é a etapa de aquisição de dados, e comumente envolve o uso de diferentes algoritmos vindos da área de processamento de imagens e reconhecimento de padrões. Esses algoritmos são aplicados junto a padrões sintéticos cujas características físicas permitem uma rápida captura e reconhecimento de

pontos de referência, conhecidos como marcadores, os quais formam parte da sua estrutura. Muitos desses padrões foram propostos inicialmente em métodos de calibração de câmera individual [16][18], e depois foram reutilizados para o processo de calibração de múltiplas câmeras. Podemos identificar 3 tipos característicos de padrão (Figura 1):

- Padrão planar.
- Padrão unidimensional.
- Padrão adimensional.

Os padrões de calibração têm como objetivo permitir a captura de pontos de referência a partir da detecção e do reconhecimento de características implícitas do seu formato. A amostra de pontos de referência extraída a partir da análise e do processamento de vistas capturadas do padrão é usada no processo de calibração tanto na etapa de calibração inicial como numa etapa final de otimização do valor dos parâmetros obtidos pela calibração inicial.

O uso de padrões contendo esses pontos de referência tem sido estudado desde os métodos desenvolvidos por Tsai [16] e Zhang [18], onde o padrão de calibração usado tem um formato planar (*padrão planar*) e contém na sua superfície pontos de referência definidos por marcadores explícitos ou implícitos desenhados e/ou inseridos na superfície do padrão.

Outros padrões propostos para calibração de sistemas estéreo e de múltiplas câmeras surgiram na última década, dentro de um grupo considerado como *padrões unidimensionais e adimensionais* comumente usados em abordagens de auto-calibração de câmera apresentadas por Borghese e Cerveri [2] e Sbovoda et al.[14]. Esses trabalhos mostram a nova tendência que é a proposta de padrões mais simples, compostos por um ou dois marcadores, que trazem como principal vantagem a fácil manipulação, movimentação e reconhecimento nas diversas vistas capturadas do padrão dentro do espaço de rastreamento.



Figura 1. Padrões planar [18], unidimensional [2] e adimensional [13].

Também podemos citar outro tipo de padrão conhecido como padrão tridimensional composto por um conjunto de dois ou três planos colocados perpendicularmente um em relação ao outro. Este tipo de padrão é pouco utilizado para a calibração de múltiplas câmeras pelos problemas de oclusão quando mais de uma câmera tenta enxergar

simultaneamente os marcadores inseridos na sua estrutura, por esse motivo ele não é considerado na nossa análise sobre calibração de múltiplas câmeras

## 2.2 Segunda etapa: Calibração Inicial

Esta etapa agrupa e descreve os diferentes métodos utilizados para calcular o valor inicial dos parâmetros intrínsecos, extrínsecos e da distorção das lentes das múltiplas câmeras que compõem o sistema. O cálculo dos parâmetros das câmeras é definido como calibração de câmera.

Em geral, as diferentes técnicas propostas na literatura para calibração de câmera podem ser agrupadas em duas categorias [5]: calibração fotogramétrica e auto-calibração.

### 2.2.1. Calibração fotogramétrica

Neste tipo de método calculam-se os parâmetros de calibração de uma câmera usando a informação fornecida por um padrão cuja geometria é conhecida com precisão no espaço tridimensional onde ele é posicionado [5]. Um exemplo clássico desse tipo de método é o proposto por Tsai [16], que usa um padrão planar montado num sistema de trilhos que movimentam o padrão com precisão em frente à câmera. Outro trabalho semelhante, porém mais flexível, é proposto por Zhang [18], que também utiliza um padrão planar movimentado livremente diante da câmera.

Com base nas informações fornecidas pelos pontos referenciais colocados sobre o padrão planar, encontra-se uma correspondência direta entre a posição 3D dos marcadores nesse padrão e os pontos 2D correspondentes das imagens capturadas do padrão.

Este tipo de método gera resultados muito precisos, mas na prática, é pouco utilizado por causa do alto custo requerido para montar um padrão planar que ocupe o maior campo de visão da câmera, além de movimentá-lo, capturá-lo e reconhecê-lo com precisão em cada vista capturada do padrão. Este método também está mais relacionado ao processo de calibração individual de câmera, e é muito utilizado especialmente para calcular os parâmetros intrínsecos e os coeficientes de distorção da lente.

### 2.2.2. Auto-calibração

Neste tipo de método, a calibração dos parâmetros intrínsecos e extrínsecos é feita a partir das informações fornecidas por pontos correspondentes corretamente identificados em imagens capturadas pela(s) câmera(s). Nesse sentido, este tipo de método não precisa de padrões complexos com uma estrutura bem conhecida, como os métodos fotogramétricos.

Os pontos correspondentes podem ser capturados usando padrões unidimensionais ou adimensionais.

No caso de usar diferentes câmeras é preciso que os pontos referenciais estejam devidamente correlacionados e identificados entre as diversas câmeras que visualizam um mesmo cenário. Pode-se afirmar que esses métodos assumem que a cena visualizada é estática, e é só baseado na informação 2D dos pontos referenciais capturados e corretamente correlacionados, é que calculamos os parâmetros de calibração das câmeras. Esse tipo de método é também muito utilizado em aplicações que realizam reconstrução 3D.

A área de pesquisa que explora essa nova estratégia de calibração de câmera e define esse tipo de método é conhecida como “estrutura a partir do movimento” (structure from motion) [6]. A base da implementação deste tipo de método se encontra na teoria da geometria epipolar, representada principalmente pelo cálculo da matriz fundamental. Este tipo de método é preferentemente usado para calibração dos parâmetros extrínsecos que relacionam a posição de uma câmera em relação a outra.

## 2.3 Terceira etapa: Otimização da calibração

Esta última etapa agrupa os diferentes métodos usados para otimizar o valor inicial dos parâmetros de calibração, calculados na etapa de calibração inicial. As soluções obtidas para os parâmetros das câmeras, tanto com o método fotogramétrico como com o de auto-calibração, são calculadas com base na minimização algébrica de um sistema de equações que modela a correspondência entre pontos 3D-2D (fotogramétrica) e pontos 2D-2D (auto-calibração). Essas soluções satisfazem uma aproximação algébrica linear para um problema que não é linear, como o processo de calibração de câmera.

A minimização algébrica pode também não levar em conta outras propriedades físicas do modelo de câmera real que influenciam o posicionamento dos pontos de referência usados no processo de calibração, como por exemplo a distorção causada pela lente física de uma câmera. Por esse motivo e pela definição de não-linearidade do processo de calibração, é requerida a aplicação de algoritmos de otimização que minimizem o erro produzido na avaliação das características físicas implícitas, que vem junto as informações fornecidas pelos pontos referenciais capturados a partir dos padrões de calibração.

Os algoritmos de minimização não-linear então tentam aprimorar a estimativa inicial obtida para os parâmetros intrínsecos, extrínsecos e coeficientes de distorção da lente das múltiplas câmeras incluídas no sistema. Para isso, esses algoritmos precisam da definição de uma função global de erro que avalie, a

cada iteração, se a nova solução estimada a partir de alguma variação da solução inicial gera melhores resultados. Muitas vezes essa função de erro global não é única e é definida por um conjunto de funções de controle especificamente criadas e relacionadas a propriedades e características físicas presentes tanto no processo de calibração como no formato definido entre os pontos referenciais capturados a partir dos padrões de calibração. Algumas das funções de controle comumente utilizadas são descritas a seguir.

### 2.3.1 Reprojção 3D – 2D

Este tipo de função é utilizada quando temos informação tanto da posição 2D na imagem como da posição 3D no espaço dos pontos referenciais capturados em nossa amostra para calibração. Com essa informação e a estimativa dos parâmetros de calibração, nós podemos reprojeter sobre o plano da imagem os pontos 3D reconstruídos, e essa reprojção deve ficar muito próxima a posição dos pontos 2D correspondentes extraídos diretamente da imagem usando algum algoritmo de processamento de imagem. O somatório da diferença entre a posição do ponto 2D extraído a partir da imagem e o calculado com base no processo de reprojção é o valor de erro determinado para esse tipo de função de controle.

### 2.3.2 Reprojção 2D – 2D

Este tipo de função é definida como a distância da linha epipolar ao ponto correspondente num sistema estéreo de câmeras. Para poder modelar essa função precisamos das informações sobre os pontos correspondentes 2D  $m_i \leftrightarrow m_i'$ , extraídos a partir de duas imagens de um mesmo cenário. Para poder relacionar estes pontos é preciso ter já calculado o valor da matriz fundamental que as relaciona.

A A relação estabelecida pela matriz fundamental permite a projeção de um ponto  $m$  numa câmera A sobre o plano da imagem de uma câmera B. A matriz fundamental  $F$  determina a relação entre as câmeras A e B. A projeção do ponto  $m$  então cria uma linha  $l'$  no plano da imagem da câmera B e sobre essa linha é que deve estar posicionado o ponto correspondente  $m'$ . No caso ideal, a distância do ponto  $m$  à linha  $l'$  deve ser zero quando o ponto  $m$  cai exatamente sobre a linha  $l'$ . No entanto, por causa do ruído e do valor inicial dos parâmetros da câmera, essa distância fica perto de zero.

Essa função é definida como uma projeção 2D-2D, na qual o erro estabelecido é medido pela distância de um ponto  $m'$  à sua linha correspondente  $l'$  e pela distância de um ponto  $m$  à sua linha correspondente  $l$ .

### 2.3.3 Distância 3D entre pontos reconstruídos

Este tipo de função é definida como a distância 3D que pode ser recuperada quando temos nossos pontos correspondentes  $m_i \leftrightarrow m_i'$  reconstruídos no espaço determinado pelo campo de visão das câmeras. Essa reconstrução 3D deve ser feita em escala semelhante a alguma unidade de medida métrica. Um exemplo de aplicação dessa função de controle é quando usamos padrões unidimensionais, cuja distancia entre os marcadores que os compõem é conhecida. O processo de reconstrução 3D dos pontos correspondentes  $m_i \leftrightarrow m_i'$  é feito comumente usando um algoritmo de triangulação como o apontado em [29][30], que usa todos os parâmetros da câmera para calcular a posição 3D de um marcador  $M$  junto as informações dadas por  $m_i$  e  $m_i'$ . Tendo as distâncias entre os marcadores reconstruídos poderemos compará-las com as distâncias reais medidas a partir dos marcadores físicos do padrão de calibração.

### 2.3.4 Função de erro global

As funções de controle acima descritas podem ser usadas em conjunto para modelar uma única função de erro global, que pode mudar dependendo do método aplicado e do padrão a ser usado na aquisição de dados, onde deve se garantir que haja informação de entrada suficiente para modelar as funções de controle antes descritas.

Uma vez determinados os dois componentes básicos para rodar um algoritmo de otimização – a amostra de entrada correspondente aos pontos referenciais e a definição das funções de controle e função de erro global – poderemos executar o algoritmo de otimização. Um dos algoritmos muito utilizados para esse processo é o descrito por Levenberg Marquart [4].

## 3. Método proposto

O método proposto pode ser definido como um método híbrido, sendo composto por um método fotogramétrico e um de auto-calibração trabalhando em conjunto. A mistura desses dois tipos de métodos de calibração básicos é feita com os seguintes propósitos:

- O uso de um método fotogramétrico permite calcular uma estimativa inicial confiável dos parâmetros intrínsecos e dos coeficientes de distorção da lente para cada câmera do sistema. Esse tipo de método é útil porque permite recuperar uma estimativa inicial dos parâmetros intrínsecos e dos coeficientes de distorção da lente de câmeras cujas informações específicas são desconhecidas.

- A utilização de um método de auto-calibração permite realizar o cálculo dos parâmetros extrínsecos que relacionam a posição de uma câmera com a de outra. A ideia por trás do uso deste tipo de método num sistema com múltiplas câmeras é realizar a calibração dos parâmetros extrínsecos de todas as possíveis combinações de duas câmeras que podem ser geradas para nosso sistema de “n” câmeras e, com esses resultados, encontrar a melhor calibração, duas a duas câmeras, que permita relacionar todas as câmeras do sistema, criando uma concatenação entre elas. A ideia de concatenação das calibrações dos diversos sistemas estéreo foi proposta no trabalho de Kurillo et al. [9], que se mostrou muito útil quando se precisou calibrar múltiplas câmeras que cobriam uma área de rastreamento muito ampla e bem distribuída.

A combinação dos dois tipos de calibração por parte do método proposto permite uma maior flexibilidade e robustez no cálculo de todos os parâmetros de calibração, especialmente quando se tem equipamentos muito variados e áreas de rastreamento amplas.

Seguindo as etapas definidas pelo framework conceitual, descreveremos a seguir os principais componentes de nosso método.

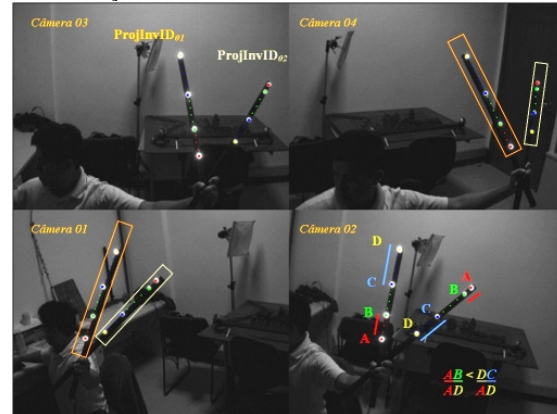
### 3.1 Aquisição de dados

O método proposto faz uso de dois tipos de padrão de calibração: um padrão planar e um padrão unidimensional, definido como **padrão invariante**.

Nesta etapa podemos ressaltar como primeira contribuição a proposta e uso de um novo tipo de padrão, o **padrão invariante**. Este padrão é formado por um conjunto de 4 marcadores distribuídos colinearmente, e traz como principal novidade uma técnica robusta desenvolvida em Loaiza et al.[10] e testada por Teixeira et al.[15] para reconhecer individualmente diferentes arranjos de 4 marcadores colineares, onde a única exigência é que os arranjos de marcadores tenham diferentes distribuições dos marcadores, isto é, as distâncias entre os marcadores de dois arranjos sejam diferentes. Para cada arranjo é gerado um identificador único definido como **ProjInvID<sub>i</sub>** (Figura 2).

Esta técnica então permite que possamos usar simultaneamente mais de um padrão invariante na cena para captura de pontos de referência usados no processo de calibração, e esta é uma vantagem não encontrada em nenhum dos outros métodos de calibração pesquisados na literatura. Cada padrão invariante é livre e independente na sua movimentação dentro da cena.

Além de permitir a captura de mais de um padrão invariante dentro da cena, este padrão traz mais informação física implícita no seu formato, se comparado ao clássico padrão 1D formado por uma vara com dois marcadores [2][17]. O padrão também traz duas vezes mais pontos de referência por vista capturada do padrão (4 pontos de referência) e duas distâncias a mais entre marcadores. Essas informações adicionais entre os marcadores serão de muita ajuda no processo otimização dos parâmetros de calibração



**Figura 2.** Captura e identificação dois padrões invariantes nas 4 câmeras do sistema.

### 3.2 Calibração inicial

Nesta etapa, por causa da característica híbrida do método, também fazemos uso de dois tipos de calibração para achar diferentes parâmetros das câmeras.

Em relação a calibração fotogramétrica utilizamos como referência o método de Zhang [18] o qual usa o padrão planar. Para sua implementação foi usada a versão fornecida pela biblioteca OpenCV [7]. Os parâmetros das câmeras recuperados são os parâmetros intrínsecos e os coeficientes de distorção das lentes.

A informação necessária para executar o método de calibração fotogramétrico é fornecida pela amostra de pontos referenciais obtidos a partir da captura e do processamento das vistas do padrão planar. Cada câmera do sistema é calibrada individualmente usando o padrão planar usando amostras contendo entre 50 e 200 vistas do padrão planar.

Em relação ao método de auto-calibração usado neste trabalho, usamos o método descrito por Dorfmueller-Ulhaas em [5]. Este processo pode ser definido como a calibração dos “m” sistemas estéreos definidos pelo número de combinação ( $C^n_2$ ) onde “n” é o número de câmeras que formam o sistema. Os parâmetros extrínsecos recuperados são a matriz de rotação para cada par de câmeras e o vetor translação. A informação necessária para a execução

do método de auto-calibração é fornecida pela amostra de pontos referenciais obtidos a partir da captura e do processamento das vistas do **padrão invariante** e da estimativa inicial dos parâmetros calculados com o método de calibração fotogramétrica (**padrão planar**).

Um processo importante no método de autocalibração é o cálculo da matriz fundamental. Usamos um algoritmo iterativo de RANSAC junto com o algoritmo de 8 pontos. A implementação deste algoritmo presente no OpenCV foi utilizada.

Uma vez calculada a matriz fundamental, o passo seguinte é a extração da matriz essencial usando os parâmetros intrínsecos calculados na calibração fotogramétrica. É a partir da matriz essencial que recuperaremos a orientação e a posição de uma câmera em relação a seu par correspondente. O processo de autocalibração é realizado em cada um dos  $m$  ( $= C^n_2$ ) sistemas de câmera em estéreo definidos para nosso sistema. Na implementação, todos os processos referentes à recuperação da posição e da orientação das câmeras foram implementados seguindo os algoritmos e os detalhes descritos em [5][6][11].

Terminado o cálculo de uma solução inicial para todos os parâmetros das câmeras, procederemos com a etapa de otimização desses parâmetros.

### 3.3 Otimização dos parâmetros

Nesta terceira etapa do processo de calibração, descrevemos como o método proposto usa toda a informação implícita relacionada às características do formato do padrão invariante para definir e modelar distintas funções de controle com as quais montaremos uma função de erro global, usada na otimização dos parâmetros das câmeras. Como foi descrito na seção 2, as funções de controle são modeladas com base na estimativa ou medição de certas características extraídas do formato do padrão, que depois são modeladas e usadas como funções matemáticas que internamente levam em conta os valores dos parâmetros das câmeras que desejamos otimizar.

Quanto mais informações pudermos extrair do formato do padrão, melhor será a modelagem matemática das características nele implícitas. Portanto, as funções de controle são criadas para quantificar e controlar o valor correto dessas características quando são testadas novas soluções para os parâmetros das câmeras geradas durante a execução do algoritmo de otimização.

Para o caso do padrão invariante podemos observar que as características físicas fornecidas pela sua própria estrutura permitem que todas as funções de controle descritas na seção 2.3 sejam usadas. Mas também podemos notar mais uma característica que

pode ser explorada e modelada como mais uma função de controle, que é a **colinearidade** dos marcadores do padrão. A característica da colinearidade também permanece invariante quando os marcadores são projetados sobre o plano da imagem da câmera. Isto quer dizer que podemos modelar duas funções de controle relacionadas à colinearidade desses pontos, uma para medir a colinearidade dos marcadores do padrão após sua reconstrução 3D, e outra para medir a colinearidade dos mesmos quando extraímos sua posição 2D sobre o plano da imagem.

Todo esse conjunto de funções de controle permite modelar uma função de erro global mais robusta, pois contém mais informações das características do padrão, o que nos ajuda a melhorar o cálculo e convergência do processo de otimização.

## 4. Ambiente de teste e resultados

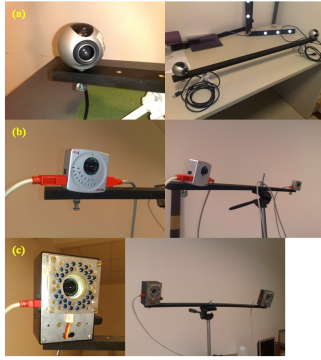
O objetivo dos testes realizados é mostrar as contribuições que nosso método pode oferecer, aproveitando especialmente as características implícitas do novo padrão invariante. Tenta-se provar que:

- A característica de colinearidade do padrão contribui na obtenção de melhores resultados quando usada na etapa de otimização dos parâmetros das câmeras. Compararemos os resultados obtidos incluindo ou não essas funções derivadas da colinearidade do padrão.
- A precisão alcançada na tarefa de recuperação de medidas 3D, após realizar a calibração das câmeras com nosso método, é melhor em comparação aos resultados obtidos com outros métodos de calibração de múltiplas câmeras, como os apresentados academicamente em [2][13][14] e [17] e suas versões comerciais [1] e [8], que serão citados como referência para comparação com nosso método.

### 4.1 Equipamento de teste

Para realizar os testes, foram montados dois sistemas formados por 4 câmeras:

- 2 câmeras web, Logitech, modelo QuickCam Pro 4000, conexão USB 2.0, captura 30 fps. (Figura 3a).
- 4 câmeras monocromáticas, Unibrain, modelo Fire-i, conexão firewire, captura 30 fps. (Figura 3b e 3c).



**Figura 3.** (a) Câmeras web Logitech, (b) e (c) Fire-I  
Os sistemas de câmeras montados foram compostos da seguinte forma:

- Sistema X: composto pelas quatro câmeras firewire (Figura 4a ).
- Sistema Y: composto por duas câmeras firewire e as duas câmeras web (Figura 4b ).



**Figura 4.** (a) Sistema X e (b) sistema Y de quatro câmeras.

Também foram montados dois padrões invariantes formados por duas varas de madeira onde foram colocadas 4 lâmpadas pequenas alinhadas colinearmente (Figura 2), e um padrão planar para a calibração fotogramétrica inicial de cada câmera.

As distâncias a que foram colocadas as lâmpadas em cada padrão invariante foram:

- Padrão 01: AB (39.01 mm) , BC (140.196 mm), CD (99.156 mm).
- Padrão 02: AB (99.65 mm) , BC (139.505 mm), CD (180.565 mm).

Para executar os testes foi usado um computador composto por um processador Intel Core 2 Duo de 2.3Ghz, com 2GB de memória e foi necessária uma placa PCI de extensão de portas firewire para prover

mais duas entradas firewire e assim poder conectar nossas quatro câmeras num só computador.

Todo o processo de captura, processamento das imagens e execução do algoritmo de calibração de múltiplas câmeras foi feito num só computador, ressaltando principalmente que os processos de captura de vídeo e processamento das imagens vindas das quatro câmeras do sistema rodaram numa resolução de imagem de 640x480 pixels mantendo uma taxa de captura constante de ~30fps em todas as câmeras. A área aproximada de área de rastreamento tem as dimensões 2 m x 2 m x 1.8 m que é um volume padrão similar a outros testes realizados em trabalhos como [2][5][13] e [17].

## 4.2 Metodologia de teste

Para testar os resultados da calibração de câmeras com o método proposto, identificamos que a maioria de métodos existentes na literatura mostra sua real contribuição quando realiza o cálculo final dos parâmetros das câmeras, na etapa de otimização dos parâmetros. Observamos que uma boa forma de comparar esses resultados é verificar o erro parcial alcançado com os parâmetros otimizados quando eles são reutilizados para medir o erro obtido para cada função de controle usada no processo de otimização. Quanto menor o erro, melhor será a estimativa dos parâmetros de calibração das câmeras.

É assim que nós propomos a seguinte metodologia de testes formularemos diferentes funções de erro global, as quais serão usadas no algoritmo de otimização para aprimorar o valor dos parâmetros das câmeras. Inicialmente criaremos funções de erro global baseadas só em funções de controle convencionais como: reprojeção 3D-2D ( $\mathcal{E}_{Reproj.3D-2D}$ ), reprojeção 2D-2D ( $\mathcal{E}_{Reproj. 2D - 2D}$ ) e reconstrução 3D dos marcadores ( $\mathcal{E}_{dist. 3D}$ ). Depois, usaremos as mesmas definições de erro global e adicionaremos em cada uma delas as funções derivadas da colinearidade, que pode ser explorada graças ao formato dos padrões invariantes. A seguir mostraremos as 8 funções de erro que formulamos para realizar os testes e as comparações:

- Função de erro 1: simula o ambiente de calibração definido pelos métodos de Borghese et al.[2] , Kurillo et al.[9] e Uematsu et al.[17].

$$\mathcal{E}_{Global} = (\mathcal{E}_{Reproj. 2D - 2D} + \mathcal{E}_{dist. 3D})$$

- Função de erro 2, versão da função impar 1 com funções de controle derivadas da colinearidade do nosso padrão.

$$\mathcal{E}_{Global} = (\mathcal{E}_{Reproj.2D-2D} + \mathcal{E}_{dist. 3D} + \mathcal{E}_{Colinearidade 3D} + \mathcal{E}_{Colinearidade 2D})$$

- Função de erro 3: simula o ambiente de calibração definido pelo método de Bouguet [3].

$$\mathcal{E}_{Global} = (\mathcal{E}_{Reproj. 3D-2D} + \mathcal{E}_{dist. 3D})$$

- Função de erro 4.

$$\mathcal{E}_{Global} = (\mathcal{E}_{Reproj. 3D-2D} + \mathcal{E}_{dist. 3D} + \mathcal{E}_{Colinearidade 3D} + \mathcal{E}_{Colinearidade 2D})$$

- Função de erro 5: simula um ambiente de calibração onde todas as funções de controle convencionais podem ser utilizadas, definindo uma função de erro global convencional robusta.

$$\mathcal{E}_{Global} = (\mathcal{E}_{Reproj. 3D-2D} + \mathcal{E}_{Reproj. 2D-2D} + \mathcal{E}_{dist. 3D})$$

- Função de erro 6.

$$\mathcal{E}_{Global} = (\mathcal{E}_{Reproj. 3D-2D} + \mathcal{E}_{Reproj. 2D-2D} + \mathcal{E}_{dist. 3D} + \mathcal{E}_{Colinearidade 3D} + \mathcal{E}_{Colinearidade 2D})$$

- Função de erro 7: simula o ambiente de calibração definido pelos métodos descritos por Svoboda et al. [14] e Pintaric et al. [13]. Uma única função de controle é modelada como função de erro global devido às poucas informações fornecidas pelo tipo de padrão de calibração adimensional usada nesse método.

$$\mathcal{E}_{Global} = (\mathcal{E}_{Reproj. 3D-2D})$$

- Função de erro 8.

$$\mathcal{E}_{Global} = (\mathcal{E}_{Reproj. 3D-2D} + \mathcal{E}_{Colinearidade 3D} + \mathcal{E}_{Colinearidade 2D})$$

Nosso objetivo com a metodologia de teste proposta é provar que realmente o novo padrão invariante e a característica de colinearidade que possui geram informação adicional que melhoram o cálculo dos parâmetros das câmeras quando é aproveitada no algoritmo de otimização.

### 4.3 Descrição de resultados

Para rodar nossos testes, o procedimento de teste foi definido da seguinte forma:

- Para cada sistema X e Y capturas-se uma amostra “A” contendo de 250 vistas do padrão invariante.
- Depois capturasse outra amostra “B” de 1000 vistas do padrão invariante.
- Com a amostra A calibra-se cada sistema de múltiplas câmeras, obtendo 8 conjuntos de parâmetros para as câmeras, uma para cada das 8 funções de erro descritas na seção 4.2.

Finalmente, com os 8 conjuntos de parâmetros das câmeras já calibrados, testa-se a precisão alcançada ao avaliar esses parâmetros na amostra B. Especialmente focaremos os resultados alcançados ao avaliar as funções convencionais encontradas na

maioria dos métodos analisados. Os resultados dos testes são apresentados nas tabelas a seguir:

**Tabela 1.** Erro médio medido na amostra B, capturada para a calibração do sistema X.

| Câmeras 1 e 2             | Função 01 | Função 02 | Função 03 | Função 04 | Função 05 | Função 06 | Função 07 | Função 08 |
|---------------------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| Reprojeção 3D-2D (pixels) | 0,403     | 0,391     | 0,444     | 0,422     | 0,402     | 0,391     | 0,369     | 0,371     |
| Distância 3D (mm)         | 1,499     | 1,472     | 1,392     | 1,455     | 1,501     | 1,478     |           |           |
| Reprojeção 2D-2D (pixels) | 0,806     | 0,781     | 0,889     | 0,844     | 0,805     | 0,782     | 0,738     | 0,7411    |
| Câmeras 2 e 3             |           |           |           |           |           |           |           |           |
| Reprojeção 3D-2D (pixels) | 0,426     | 0,422     | 0,439     | 0,4372    | 0,421     | 0,424     | 0,397     | 0,394     |
| Distância 3D (mm)         | 1,691     | 1,664     | 1,719     | 1,713     | 1,677     | 1,666     |           |           |
| Reprojeção 2D-2D (pixels) | 0,852     | 0,844     | 0,877     | 0,874     | 0,842     | 0,847     | 0,793     | 0,788     |
| Câmeras 3 e 4             |           |           |           |           |           |           |           |           |
| Reprojeção 3D-2D (pixels) | 0,133     | 0,132     | 0,182     | 0,177     | 0,128     | 0,127     | 0,14      | 0,135     |
| Distância 3D (mm)         | 0,794     | 0,791     | 0,785     | 0,772     | 0,801     | 0,794     |           |           |
| Reprojeção 2D-2D (pixels) | 0,267     | 0,264     | 0,364     | 0,355     | 0,256     | 0,254     | 0,28      | 0,27      |

**Tabela 2.** Erro médio medido na amostra B, capturada para a calibração do sistema Y.

| Câmeras 1 e 2             | Função 01 | Função 02 | Função 03 | Função 04 | Função 05 | Função 06 | Função 07 | Função 08 |
|---------------------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| Reprojeção 3D-2D (pixels) | 0,716     | 0,691     | 0,632     | 0,599     | 0,765     | 0,762     | 0,581     | 0,575     |
| Distância 3D (mm)         | 1,914     | 1,694     | 1,262     | 1,174     | 2,164     | 2,162     |           |           |
| Reprojeção 2D-2D (pixels) | 1,432     | 1,382     | 1,263     | 1,199     | 1,529     | 1,524     | 1,16      | 1,15      |
| Câmeras 2 e 3             |           |           |           |           |           |           |           |           |
| Reprojeção 3D-2D (pixels) | 1,5514    | 1,5506    | 1,556     | 1,559     | 1,557     | 1,556     | 1,552     | 1,551     |
| Distância 3D (mm)         | 1,816     | 1,809     | 1,833     | 1,83      | 1,795     | 1,787     |           |           |
| Reprojeção 2D-2D (pixels) | 3,099     | 3,098     | 3,111     | 3,114     | 3,112     | 3,109     | 3,098     | 3,097     |
| Câmeras 3 e 4             |           |           |           |           |           |           |           |           |
| Reprojeção 3D-2D (pixels) | 0,5195    | 0,5194    | 0,557     | 0,546     | 0,518     | 0,517     | 0,513     | 0,512     |
| Distância 3D (mm)         | 1,095     | 1,093     | 1,115     | 1,068     | 1,116     | 1,092     |           |           |
| Reprojeção 2D-2D (pixels) | 1,0389    | 1,0387    | 1,114     | 1,093     | 1,035     | 1,034     | 1,025     | 1,024     |

Nas tabelas 1 e 2 mostramos os resultados do erro médio obtidos ao avaliar cada uma das funções de controle usadas para comparação, já nas tabelas 3 e 4 são apresentados uma comparação percentual do erro médio das funções de número par, que levaram em conta a colinearidade, e as funções de número ímpar que não usaram a informação adicional sobre colinearidade do padrão.

**Tabela 3.** Ganho percentual ao analisar os resultados apresentados na tabela 1.

| Câmeras 1 e 2        | Ganho % função 02 para 01 | Ganho % função 04 para 03 | Ganho % função 06 para 05 | Ganho % função 08 para 07 |
|----------------------|---------------------------|---------------------------|---------------------------|---------------------------|
| Reprojeção 3D-2D (%) | 2,98                      | 4,95                      | 2,74                      | -0,54                     |
| Distância 3D (%)     | 1,8                       | -4,53                     | 1,53                      |                           |
| Reprojeção 2D-2D (%) | 3,1                       | 5,06                      | 2,86                      | -0,42                     |
| Total acumulado (%)  | 7,88                      | 5,48                      | 7,13                      | -0,96                     |
| Câmeras 2 e 3        |                           |                           |                           |                           |
| Reprojeção 3D-2D (%) | 0,92                      | 0,41                      | -0,71                     | 0,76                      |
| Distância 3D (%)     | 1,59                      | 0,35                      | 0,66                      |                           |
| Reprojeção 2D-2D (%) | 0,94                      | 0,34                      | -0,59                     | 0,63                      |
| Total acumulado (%)  | 3,45                      | 1,1                       | -0,64                     | 1,39                      |
| Câmeras 3 e 4        |                           |                           |                           |                           |
| Reprojeção 3D-2D (%) | 0,75                      | 2,74                      | 0,78                      | 3,57                      |
| Distância 3D (%)     | 0,38                      | 1,66                      | 0,87                      |                           |
| Reprojeção 2D-2D (%) | 1,12                      | 2,47                      | 0,78                      | 3,57                      |
| Total acumulado (%)  | 2,25                      | 6,87                      | 2,43                      | 7,14                      |



**Tabela 4.** Ganho percentual ao analisar os resultados apresentados na tabela 2.

|                      | Ganho %<br>função 02 para<br>01 | Ganho %<br>função 04 para<br>03 | Ganho %<br>função 06 para<br>05 | Ganho %<br>função 08 para<br>07 |
|----------------------|---------------------------------|---------------------------------|---------------------------------|---------------------------------|
| <b>Câmeras 1 e 2</b> |                                 |                                 |                                 |                                 |
| Reprojeção 3D-2D (%) | 3,49                            | 5,19                            | 0,39                            | 1,03                            |
| Distância 3D (%)     | 11,49                           | 6,97                            | 0,09                            |                                 |
| Reprojeção 2D-2D (%) | 3,49                            | 5,07                            | 0,33                            | 0,86                            |
| Total acumulado (%)  | 18,47                           | 17,23                           | 0,81                            | 1,89                            |
| <b>Câmeras 2 e 3</b> |                                 |                                 |                                 |                                 |
| Reprojeção 3D-2D (%) | 0,05                            | -0,19                           | 0,07                            | 0,06                            |
| Distância 3D (%)     | 0,39                            | 0,16                            | 0,45                            |                                 |
| Reprojeção 2D-2D (%) | 0,03                            | -0,09                           | 0,09                            | 0,03                            |
| Total acumulado (%)  | 0,47                            | -0,12                           | 0,61                            | 0,09                            |
| <b>Câmeras 3 e 4</b> |                                 |                                 |                                 |                                 |
| Reprojeção 3D-2D (%) | 0,02                            | 1,97                            | 0,19                            | 0,02                            |
| Distância 3D (%)     | 0,18                            | 4,22                            | 2,15                            |                                 |
| Reprojeção 2D-2D (%) | 0,02                            | 1,89                            | 0,09                            | 0,09                            |
| Total acumulado (%)  | 0,22                            | 8,08                            | 2,43                            | 0,11                            |

Como pode ser observado nas tabelas 3 e 4, na maioria das vezes as funções pares conseguem obter um erro médio menor em relação às ímpares. Houve casos em que se alcançou um ganho de 18%. O ganho percentual médio ficou em 3,9 %, o que já representa uma melhoria dos resultados obtidos quando usados os parâmetros das câmeras obtidos da calibração com funções que levaram em conta as funções derivadas da colinearidade do nosso padrão. É importante ressaltar também que nosso método não é infalível e, como também se observa em nossas tabelas, houve alguns casos em que as funções pares não tiveram ganho, porém a perda no pior caso não foi maior que 0,96%. Com os resultados mostrados nas tabelas acima podemos afirmar que a característica de colinearidade ajudou a melhorar a calibração dos parâmetros das câmeras, provando a contribuição que nosso método propõe com o uso dos novos padrões invariantes.

Outro resultado importante é o valor do erro médio alcançado em relação à recuperação das distâncias 3D entre os marcadores. Ele é um valor muito importante medido para comparar sistema de rastreamento comerciais orientados a aplicações médicas e de captura de movimento. A precisão obtida neste tipo de medida nos sistemas comerciais não deve passar os 5 mm para áreas de rastreamento semelhantes às usadas em nossos testes (2 m × 2 m × 2 m). Com base nos valores obtidos e mostrados para este tipo de erro em nossas tabelas 1 e 2, podemos afirmar que o maior erro obtido foi de 2,162 mm e o menor foi de 0,772mm. O erro médio geral analisando tanto os sistemas X e Y foi de 1.43 ± 0,7mm. Esse valor é muito menor que o máximo permitido de 5 mm.

A seguir mostraremos a comparação da precisão do erro alcançado por outros métodos de calibração proposto em artigos acadêmicos, assim como por especificações técnicas recolhidas de alguns produtos comerciais (Tabela 5). A comparação pode ser um pouco superficial, devido aos diferentes equipamentos que podem ter sido utilizados para obter essas medidas. Porém, é importante ressaltar

que em todos os trabalhos citados, os autores afirmam que usam hardware composto por câmeras de uso semi-industrial junto a algum equipamento de sincronismo eletrônico que garante uma correta captura de imagens simultaneamente entre todas as câmeras. Em nosso caso, os resultados foram obtidos usando câmeras sem sincronismo nenhum, e no caso das câmeras web utilizadas, elas estariam classificadas como câmeras de baixa qualidade, mas mesmo assim foi possível alcançar uma boa precisão milimétrica com as mesmas.

O padrão invariante trouxe robustez ao processo de aquisição de dados e nova informação explorada pelo processo de otimização dos parâmetros das câmeras. A sua estrutura simples e compacta ofereceu vantagens em relação a outros padrões semelhantes, mas a técnica desenvolvida para sua correta e acertada identificação, faz dele uma opção mais robusta para a modelagem de um padrão de calibração.

**Tabela 5.** Erro médio da medição da distância entre marcadores.

| Método apresentado por                             | Valor do erro médio em (mm) |
|--|-----------------------------|
| <b>Acadêmicos</b>                                  |                             |
| Borghese et al. [4] (2000)                         | 2                           |
| Pintaric et al. [28] (2007)                        | 5,3                         |
| Uematsu et al. [37] (2007)                         | 2                           |
| <b>Obtido com o método proposto neste trabalho</b> | <b>1,5</b>                  |
| <b>Produtos comerciais</b>                         |                             |
| A.R.T. GmbH [1] (2009)                             | 2                           |
| IOTracker [13] (2009)                              | 5                           |
| VICON [39] (2009)                                  | 0,1                         |

Algumas outras vantagens qualitativas do padrão invariante em relação aos padrões unidimensional e adimensional podem ser resumidas da seguinte forma:

- Maior número de marcadores por vista do padrão capturado, 4 pontos de referência em comparação com 2, no caso do padrão unidimensional, e 1, no caso do adimensional.
- Capacidade de usar mais de um padrão invariante simultaneamente para a captura de amostras de pontos referenciais.
- Maior robustez em relação à identificação grupal e individual dos padrões e os marcadores que o compõem, com base na técnica de reconhecimento desenvolvida [10].

A principal desvantagem do padrão invariante é derivada de problemas de oclusão, isso quer dizer que se algum marcador do padrão é ocluído, o resto dos marcadores não poderá ser identificado. Porém, esta desvantagem está presente também nos outros tipos de padrões: unidimensional e adimensional.

## 5. Conclusões e trabalhos futuros

Neste artigo, foi proposta a modelagem e a implementação de um novo método de calibração de múltiplas câmeras que traz como novidade o uso de um novo tipo de padrão de calibração chamado de padrão invariante. As contribuições do método proposto estão diretamente relacionadas à robustez na identificação dos marcadores que compõem o novo padrão – especialmente aproveitado na etapa de aquisição de dados do nosso método – e na exploração da característica de colinearidade que o padrão invariante possui para a obtenção de melhores resultados no cálculo dos parâmetros das câmeras, realizado na etapa de otimização dos parâmetros de calibração. Testes provaram que a inserção da característica de colinearidade gerou algum tipo de melhora. Outro resultado importante foi o alto grau de precisão milimétrica conseguido com a calibração realizada com nosso método, obtendo um erro médio de 1.5mm, que é bem abaixo do limite de 5mm proposto nos sistemas comerciais de rastreamento óptico. É bom ressaltar que o erro médio se manteve nos dois sistemas de quatro câmeras X e Y os quais foram montados com câmeras de baixo custo que não têm nenhum tipo de sincronismo eletrônico entre elas, e que estão abaixo da média das câmeras usadas pelos sistemas comerciais.

Com relação ao novo padrão de calibração definido como padrão invariante, podemos ressaltar que este padrão possui as mesmas vantagens oferecidas pelos padrões unidimensional e adimensional, comumente usados em diferentes métodos de calibração. Porém, o padrão invariante oferece maior robustez no processo de identificação grupal e individual dos marcadores que o compõem, porque existe uma técnica específica para o reconhecimento e a identificação deste tipo de padrão [10].

Como trabalhos futuros para o método apresentado neste artigo, propomos a sua utilização e teste em aplicações onde a calibração de múltiplas câmeras representa o coração do funcionamento do sistema, como por exemplo: sistemas de rastreamento óptico, sistema de captura de movimento, scanner 3D e aplicações de treinamento cirúrgico. Finalmente, gostaríamos de aprofundar e aprimorar o processo de extração e identificação dos marcadores do padrão, a partir das imagens capturadas nas diferentes câmeras. Um exemplo deste aprimoramento foi apresentado no último ano pelo sistema comercial OptiTrack [12], onde se reportaram ganho de precisão derivado da aplicação desta nova técnica na extração da posição 2D dos marcadores na imagem

## 6. Bibliografia

- [1] ADVANCED REAL TIME TRACKING GmbH, <http://www.ar-tracking.de/>, Fevereiro, 2010.
- [2] Borghese A. and Cerveri P., Calibrating a video camera pair with a rigid bar. *Pattern Recognition*, Vol 33. pp. 81-95, 2000.
- [3] Bouguet J. Y., Camera Calibration Toolbox for Matlab, [http://www.vision.caltech.edu/bouguetj/calib\\_doc/htmls/example5.html](http://www.vision.caltech.edu/bouguetj/calib_doc/htmls/example5.html), Fevereiro, 2010.
- [4] Devernay F., C/C++ Minpack, <http://devernay.free.fr/hacks/cminpack.html>, Fevereiro, 2010.
- [5] Dorfmueller, K., Wirth, H., Real-Time Hand and Head Tracking for Virtual Environments Using Infrared Beacons, *Proceedings of the International Workshop on Modelling and Motion Capture Techniques for Virtual Environments*, pp. 113 – 127, London, UK, 1998.
- [6] Hartley R. and Zisserman A. *Multiple View Geometry in Computer Vision*. Cambridge University Press, Cambridge, 2000.
- [7] Intel Open Source Computer Vision Library, Intel Research, <http://sourceforge.net/projects/opencvlibrary/>, Fevereiro, 2010.
- [8] IOTRACKER ADVANCED OPTICAL MOTION, <http://www.iotracker.com/>, Fevereiro, 2010.
- [9] Kurillo G., Li Z., Bajcsy R., Wide-Area External Multi-Camera Calibration Using Vision Graphs and Virtual Calibration Object, In *Proceedings of 2nd ACM/IEEE International Conference on Distributed Smart Cameras (ICDSC '08)*, pp. 1- 9, Stanford, CA, 2008.
- [10] Loaiza M., Raposo B. A., Gattass M., A Novel Optical Tracking Algorithm for Point-Based Projective Invariant Marker Patterns, *3rd International Symposium on Visual Computing – ISVC 2007, EUA. Advances in Visual Computing - Lecture Notes in Computer Science*. Vol. 4841, pp. 160-169. Springer-Verlag, 2007.
- [11] Ma, Y., Soatto, S., Kosecká, J., Sastry, S.S. *An Invitation to 3-D Vision: From Images to Geometric Models*, Springer-Verlag, 2005.
- [12] Optical Motion Capture Solution - OptiTrack, <http://www.naturalpoint.com/optitrack/>, Fevereiro, 2010.
- [13] Pintaric T. and Kaufmann H., Affordable Infrared-Optical Pose Tracking for Virtual and Augmented Reality, *IEEE VR Workshop on Trends and Issues in Tracking for Virtual Environments*, Charlotte, NC (USA), 2007.
- [14] Svoboda T., Martinec D. and Pajdla T., A Convenient Multi-Camera Self-Calibration for Virtual Environments, *PRESENCE: Teleoperators and Virtual Environments*, MIT Press, Vol 14, Issue 4, pp. 407 - 422, August 2005.
- [15] Teixeira, L., Loaiza, M., Raposo, A. B., Gattass, M. *Augmented Reality Using Projective Invariant Patterns*. 4th International Symposium on Visual Computing – ISVC 2008 (*Advances in Visual Computing - Lecture Notes in Computer Science*, vol. 5358), p.520-529. Springer-Verlag, 2008.
- [16] Tsai, R.Y., Metrology Using Off-the-Shelf TV Cameras and Lenses, *IEEE Journal of Robotics and Automation*, Vol. 3, No. 4, pp. 323-344, August 1987.
- [17] Uematsu Y., Teshima T., Saito H., Honghua C., D-Calib: Calibration Software for Multiple Cameras System, *14th International Conference on Image Analysis and Processing (ICIAP 2007)*, 285-290, 2007.
- [18] Zhang Z., A Flexible New Technique For Camera Calibration, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(11):1330-1334, 2000.