

Calibração Automática de Múltiplos Marcadores em Ambiente de Realidade Aumentada

Felipe A. Caetano¹, Rodrigo L. S. Silva¹

¹Instituto de Ciências Exatas– Universidade Federal de Juiz de Fora (UFJF)
Rua José Lourenço Kelmer, s/n - Campus Universitário – 36036-900 – Juiz de Fora– MG – Brasil

felipe.caetano, rodigoluis@ice.ufjf.br

Área de pesquisa: Computação Gráfica.

Ano de Ingresso: 2012

Abstract. *This work's purpose is to develop an approach for Augmented Reality systems using multiple markers for redundancy, in order to maintain stability and alignment of the augmentation even when some markers are not fully visible. The proposed approach allows to build such scenario with no needed calibration step of the relations of rotation and distance between the markers, usually adopted in the pre-execution of these systems. The method keeps the relations between the visible markers and a base marker which is responsible for dictating the transformation of the virtual object projected.*

Resumo. *Este trabalho tem como objetivo desenvolver uma abordagem para sistemas de Realidade Aumentada que utilizam múltiplos marcadores para redundância de informação, visando manter a estabilidade e alinhamento da projeção mesmo quando haja oclusão de alguns marcadores. A abordagem proposta permite que tal cenário seja construído sem que seja necessária a etapa de calibragem das relações de rotação e translação entre os marcadores, normalmente adotada na pré-execução deste tipo de sistema. O método desenvolvido guarda as relações entre os marcadores visíveis e um marcador base, que é responsável por ditar as transformações do objeto virtual projetado.*

Palavras-chave: *Realidade Aumentada, Calibração, Automática, Marcadores.*

1. Introdução

Na área da Realidade Aumentada (RA) um dos principais problemas de se usar marcadores óticos está no fato de que a oclusão de um marcador, mesmo que de forma parcial, pode fazer com que o sistema deixe de reconhecer com exatidão qual o alinhamento do objeto virtual projetado. Uma técnica comumente usada para contornar este tipo de problema é a combinação de múltiplos marcadores gerando uma única projeção, ou seja, quando um dos marcadores não está mais inteiramente visível, outro marcador, este visível, e que tem as suas coordenadas conhecidas, é usado como referência para a projeção. O problema deste tipo de técnica é que para obter este efeito, todos os marcadores devem estar estáticos entre si e precisam ter sua transformação em relação a um ponto global fixo no espaço conhecido antes da execução, etapa conhecida como calibração.

Para a calibração, é necessário definir um plano no espaço e calcular, para cada marcador utilizado, qual a translação e rotação em \mathbb{R}^3 que este marcador está em relação à origem do plano definido. Para facilitar a tarefa, é comum adotar um dos marcadores como referência para os demais, eliminando assim a calibração de um marcador na cena, como sugerido em [Wang et al. 2010]. A calibração usando métodos manuais é bastante efetiva quando se está lidando com um conjunto de marcadores que têm as transformações previamente conhecidas, por exemplo, quando impressos numa mesma folha plana ou utilizando uma base que têm seus ângulos conhecidos com precisão. Quando nenhuma dessas situações é encontrada, a calibração pode se tornar extremamente exaustiva por exigir um nível de detalhe que nem sempre é possível de se obter utilizando ferramentas manuais como réguas e transferidores, ou mesmo medidores precisos como trenas à laser ou ultrassônicas. A situação se agrava mais ainda quando a distância entre os marcadores é muito grande, pois, mesmo o menor dos erros na calibração resulta em objetos virtuais projetados em lugares bastante diferentes. A obtenção dessas transformações de forma automática, em tempo de execução e sem necessitar da interação do usuário será o foco deste trabalho.

2. Trabalhos relacionados

[Siltanen et al. 2007] propõem um método semelhante ao deste trabalho para utilização de múltiplos marcadores sem a calibração. Na execução, um marcador central serve de referência para todos os outros. O cálculo da relação entre os marcadores se dá de maneira semelhante, porém como o marcador base é sempre o mesmo, o autor propõe o uso de uma estrutura de grafos para interconectar marcadores que não são visíveis na cena ao mesmo tempo. Desta forma, se houver um caminho no grafo que os conecte através de outros marcadores, a relação entre ambos pode ser conhecida, sabendo as relações de cada aresta que conecta os marcadores do caminho. Um lado negativo deste tipo de abordagem é que em determinadas situações, a quantidade necessária de cálculo para se chegar ao resultado final pode ser proibitiva. Quando um marcador é conectado com a base por vários outros marcadores, a relação entre cada um dos marcadores, ligados par a par, devem ser conhecidas e multiplicadas uma com a outra, até que se chegue no resultado final, somando a isso o custo de se encontrar o menor caminho dentro do grafo. A abordagem proposta aqui permite que o marcador base varie de acordo com o campo de visão da câmera, de forma que o próximo marcador a ser eleito como base, com certeza era visível pela base anterior, fazendo com que descobrir a transformação entre eles seja algo direto. O autor afirma que o resultado final exibido é bastante satisfatório.

[Kotake et al. 2004] propõem um método de calibragem combinando um conjunto de marcadores planos com a técnica de ajustamento de feixes perspectivos (*bundle adjustment*), porém, apesar do método não exigir que o usuário meça as distâncias e as rotações dos marcadores, ele exige que se tenha algum conhecimento prévio sobre seu arranjo, por exemplo, quando os marcadores estão em um mesmo plano ou quando estão em planos diferentes mas são co-normais. Os resultados apresentados pelos autores são bastante satisfatórios.

[Uematsu and Saito 2005] desenvolveram um método que combina os sistema de coordenadas independentes de todos os marcadores em um espaço projetivo único, e a partir desse, consegue calcular as relações entre cada um dos marcadores. Para construir o espaço projetivo, o sistema captura duas imagens aleatórias de posições diferentes do sistema em uma pré-etapa que, segundo os autores, dura cerca de 30 segundos. Ambas as imagens devem ser tiradas com todos os marcadores visíveis para que o espaço possa ser projetado. Isso pode ser um impedimento, já que em alguns casos, nem todos os marcadores são visíveis por duas tomadas de cena ou mesmo por uma única tomada. Além disso, este processo pode ser um inconveniente, pois não é desejável esperar o sistema executar a etapa de pré-execução.

[Baratoff et al. 2002] propuseram um método semi-automático de calibração, baseado também em imagens previamente tiradas e ajustamento de feixes perspectivos. O método funciona de maneira semelhante ao supracitado, transformando as coordenadas de cada marcador em uma só coordenada global, e então, calculando a relação entre cada um deles. O autor obteve experimentalmente um tempo de calibração igual a aproximadamente um segundo, onde seria possível avaliar até quarenta fotos e escolher o melhor par para realizar os cálculos. [Klopschitz and Schmalstieg 2007] revela um problema no método anterior que já foi citado aqui: ele só é eficiente quando todos os marcadores se encontram razoavelmente próximos, a ponto de ser possível que todos ou quase todos saiam na mesma foto. Por isso, o autor propôs um método semelhante para grandes áreas, onde uma câmera navega pelo espaço capturando os marcadores e características naturais como quinas presentes e guarda as relações em uma base. [Wang et al. 2010] também trabalhou de forma semelhante a [Baratoff et al. 2002], porém, ao invés de computar as transformações entre os marcadores até a base de forma "gulosa", utilizou o algoritmo de *Floyd-Warshall*, usando como peso entre as arestas um valor de erro pré-calculado em função do ângulo e distância dos marcadores.

Todos os trabalhos encontrados que propõem um sistema semelhante ao desenvolvido aqui têm em comum o fato de usarem métodos complexos ou etapas de pré-execução. Cálculos mais simples aumentam a performance da aplicação e são mais facilmente executáveis em computadores mais lentos ou em sistemas embarcados como *smartphones* ou *tablets*.

3. Método proposto

O modelo proposto se baseia na ideia de que é possível extrair as relações entre dois marcadores distintos em tempo de execução. Tendo as relações previamente armazenadas em algum momento onde ambos marcadores são visíveis ao mesmo tempo, é possível inferir a posição de um marcador invisível na cena, caso o outro continue visível. Com as relações de todos para todos guardadas, o sistema tem a capacidade de manter as projeções

estáveis com apenas um dos marcadores visível.

Dado que dois marcadores A e B estão presentes na cena e são detectados, e considerando que M_a e M_b simbolizam, respectivamente, as matrizes de transformação dos marcadores A e B, é possível calcular uma matriz M_{ba} que diz a relação entre eles usando:

$$M_b \times M_{ba} = M_a \quad (1)$$

$$M_{ba} = M_b^{-1} \times M_a \quad (2)$$

Para construir um sistema que aceite vários marcadores ao invés de apenas dois, deve-se antes de tudo definir qual será a política de substituição do marcador base. Neste trabalho, será chamado de marcador base aquele marcador que vai controlar a cena no momento, ou seja, toda a estimativa de pose e posição do objeto virtual projetado será calculada em relação à este marcador, usando também os dados das relações previamente registradas. É natural de se esperar que, obrigatoriamente, o marcador base deve estar visível e a sua escolha deve ser feita de maneira cuidadosa, para que o resultado final seja o melhor possível. A cada novo quadro do vídeo, a qualidade de cada marcador visível é calculada segundo um critério, a fim de definir qual deles controlará a cena.

4. Situação atual e trabalhos futuros

O sistema vem sendo desenvolvido sobre a biblioteca Artoolkit [Kato 2005] para rastrear marcadores artificiais. Sistemas baseados em rastreamento óptico apesar de serem eficientes, nem sempre são eficazes. No sistema proposto, em execuções mais longas com grande variação do posicionamento da câmera, há um acúmulo progressivo de erro que pode deixar o sistema de rastreamento impreciso.

Portanto, uma das metas para trabalhos futuros é melhorar a estabilidade do sistema, fazendo com que ele opere com pouco ou nenhum erro, de forma que seja possível a sua utilização durante longos períodos e em situações adversas, sem prejuízo na qualidade de execução. Isso pode ser feito tanto adotando técnicas estatísticas para minimizar o ruído adicionado às matrizes de relação entre os marcadores, quanto desenvolvendo um sistema de rastreamento próprio, voltado para este tipo de situação, em detrimento do rastreamento feito pelo Artoolkit, que é voltado para aplicações gerais. Métodos de correção de erro também podem ser aplicados para evitar que o acúmulo deste impeça maiores tempos de execução.

Outro ponto que pode ser explorado é a extração de características reais da cena para o rastreamento. Objetos reais planares podem ser usados na Realidade Aumentada como rastreadores naturais com eficiência, melhorando a usabilidade do sistema de forma geral, uma vez que o usuário não é mais obrigado a imprimir os marcadores artificiais para utiliza-lo. Alguns trabalhos [Lee and Höllerer 2009][Comport et al. 2003][Wagner et al. 2010] sugerem o uso de algoritmos de extração de características de imagens como SIFT[Lowe 2004], SURF[Bay et al. 2008] e Ferns[Özuysal et al. 2007] para tais fins e têm obtido bons resultados, tanto em qualidade de rastreamento quanto em performance.

A etapa final no desenvolvimento do projeto seria a proposta de uma aplicação prática onde este possa ser utilizado. Uma ideia inicial seria a construção de um método de simulação de Realidade Virtual imersiva usando marcadores de Realidade Aumentada. Seu funcionamento seria semelhante ao proposto em [Cruz-Neira et al. 1992], onde o usuário tem a sensação de estar imerso em um ambiente de Realidade Virtual, porém, ao invés de cercá-lo com projetores e rastrear a sua posição usando dispositivos magnéticos, ultrassônicos, ou mesmo componentes eletrônicos como acelerômetros e giroscópios, muitas vezes caros, isso seria feito com marcadores de RA. Para tal, o usuário deve usar um HMD (*Head Mounted Display*) com uma câmera acoplada e, baseado no rastreamento dos marcadores, sua posição pode ser estimada. A calibração automática desses marcadores seria um ponto crucial para que a aplicação pudesse ter usabilidade, uma vez que os marcadores não precisariam estar dispostos em uma configuração milimetricamente conhecida, e uma boa sensação de imersão, visto que os marcadores não precisam obrigatoriamente estar dispostos em um plano mas sim, cercando todo o ambiente onde o usuário se encontra.

Referências

- Baratoff, G., Neubeck, A., and Regenbrecht, H. (2002). Interactive multi-marker calibration for augmented reality applications. In *Proceedings of the 1st International Symposium on Mixed and Augmented Reality*, ISMAR '02, pages 107–, Washington, DC, USA. IEEE Computer Society.
- Bay, H., Ess, A., Tuytelaars, T., and Van Gool, L. (2008). Speeded-up robust features (surf). *Comput. Vis. Image Underst.*, 110(3):346–359.
- Comport, A. I., Marchand, E., and Chaumette, F. (2003). A real-time tracker for markerless augmented reality. In *Proceedings of the 2nd IEEE/ACM International Symposium on Mixed and Augmented Reality*, ISMAR '03, pages 36–, Washington, DC, USA. IEEE Computer Society.
- Cruz-Neira, C., Sandin, D. J., DeFanti, T. A., Kenyon, R. V., and Hart, J. C. (1992). The cave: audio visual experience automatic virtual environment. *Commun. ACM*, 35(6):64–72.
- Kato, H. (2005). *ARToolKit 2.33 Documentation (Alpha Version)*. Human Interface Technology Laboratory, University of Washington, <http://www.hitl.washington.edu/artoolkit/documentation/>. [Online; acessado em 21-Outubro-2011].
- Klopschitz, M. and Schmalstieg, D. (2007). Automatic reconstruction of wide-area fiducial marker models. In *Proceedings of the 2007 6th IEEE and ACM International Symposium on Mixed and Augmented Reality*, ISMAR '07, pages 1–4, Washington, DC, USA. IEEE Computer Society.
- Kotake, D., Uchiyama, S., and Yamamoto, H. (2004). A marker calibration method utilizing a priori knowledge on marker arrangement. In *Proceedings of the 3rd IEEE/ACM International Symposium on Mixed and Augmented Reality*, ISMAR '04, pages 89–98, Washington, DC, USA. IEEE Computer Society.

- Lee, T. and Höllerer, T. (2009). Multithreaded hybrid feature tracking for marker-less augmented reality. *IEEE Transactions on Visualization and Computer Graphics*, 15(3):355–368.
- Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vision*, 60(2):91–110.
- Siltanen, Hakkarainen, and Honkamaa (2007). Automatic marker field calibration. In *Virtual Reality International Conference (VRIC2007)*, pages 261–267.
- Uematsu, Y. and Saito, H. (2005). Ar registration by merging multiple planar markers at arbitrary positions and poses via projective space. In *Proceedings of the 2005 international conference on Augmented tele-existence, ICAT '05*, pages 48–55, New York, NY, USA. ACM.
- Wagner, D., Reitmayr, G., Mulloni, A., Drummond, T., and Schmalstieg, D. (2010). Real-time detection and tracking for augmented reality on mobile phones. *IEEE Transactions on Visualization and Computer Graphics*, 16(3):355–368.
- Wang, L., Springer, M., Heibel, T. H., and Navab, N. (2010). Floyd-warshall all-pair shortest path for accurate multi-marker calibration. In *ISMAR*, pages 277–278. IEEE.
- Özuysal, M., Fua, P., and Lepetit, V. (2007). Fast keypoint recognition in ten lines of code. In *CVPR*. IEEE Computer Society.