

به نام خدا

بررسی مقاله

دنبال کردن بلادرنگ دوربین توسط نشانه‌های زمین‌های ورزشی

یاسر سوری

دانشگاه صنعتی شریف

souri@ce.sharif.edu

۲۸ شهریور ۱۳۹۱

چکیده

در این مقاله روشی برای محاسبه‌ی مکان، جهت و فاصله کانونی دوربین به صورت بلادرنگ بیان می‌شود. یکی از ویژگی‌های بدیع این مقاله استفاده از چندین تصویر برای بالا بردن دقت در مرحله‌ی تعیین مکان دوربین است. همچنین روشی خودکار برای آغاز پروسه‌ی دنبال کردن دوربین بیان می‌شود. ضمناً در تعیین مکان دوربین، نیاز به تعیین مقادیری به صورت دستی است. این موضوع از نقاط ضعف اصلی این مقاله است.

۱ فرض‌های مسئله

در این بخش بعضی از فرضیات مسئله‌ای که قرار است در این مقاله حل شود مورد بررسی قرار می‌گیرد.

- فرض شده است که، پارامترهای pan، tilt و zoom به صورت بلادرنگ محاسبه می‌شوند.
- همچنین، پارامترهای مکان دوربین، roll و lens distortion در طول زمان تغییر نمی‌کنند. دقت کنید که در این مقاله roll صفر در نظر گرفته نشده است.
- در این مقاله principal point در وسط تصویر در نظر گرفته شده است. ضمناً از lens distortion این مسئله صرف نظر شده است.
- برای ورزش‌های مثل فوتبال که طول زمین ممکن است تغییر بکند، لازم است که به طور دقیق زمین اندازه‌گیری شود.
- برخی از این فرض‌ها، باعث کاهش دقت واسنجی خواهد شد. البته با این کاهش هنوز دقت واسنجی به اندازه‌ای هست که برای ورزشی مثل فوتبال مناسب باشد.

۱.۱ ایده گرفتن از مقالات دیگر

در این مقاله از دو مقاله‌ی دیگر ایده گرفته شده است و بهبودهایی روی آن‌ها انجام شده است. این مسئله بسیار برای من جالب بود. در ادامه این ایده‌گیری را توضیح خواهم داد. در این مقاله دو بخش اصلی وجود دارد. اول مقدار دهی اولیه^۱ که از ایده‌ی مقاله [۲] استفاده شده است. دوم بخش دنبال کردن است که از ایده‌ی مقاله [۱] استفاده شده است. در بخش مقدار دهی اولیه و آغاز فرآیند دنبال کردن در [۲] از یک روش جستجوی خسته کننده^۲ استفاده کرده است. این جستجو کند بوده و برای عملکرد بلادرنگ مناسب نمی‌باشد، در این مقاله بهبودی روی آن داده شده است که سرعت آن را تا ۱ ثانیه بالا می‌برد. در بخش دنبال کردن مقاله‌ی [۱] روشی بر اساس دنبال کردن خط مطرح شده است که در این مقاله تشخیص خط آن بهبود داده شده است. این نکته که ایده‌ی این مقاله کاملاً در چند مقاله‌ی دیگر

^۱ initialization

^۲ exhaustive

بوده است برای بنده جالب بود. در این مقاله عملاً با ایده‌های خوب و به جا، توانسته یک سیستم مناسب برای واسنجی دوربین ارائه کند که توانسته در دو سامانه‌ی تجاری مورد استفاده قرار بگیرد.

۲ کلیت روش

این مقاله دارای دو الگوریتم متفاوت است:

- جستجو: خروجی این قسمت، پیدا کردن ماتریس دوران و همچنین فاصله‌ی کانونی است.
- دنبال کردن: این قسمت یک مسئله‌ی بهینه سازی است که می‌توان، آن را برای پیدا کردن تمام پارامترها یا بخشی از آن‌ها حل کرد.

همچنین این مقاله دارای سه قسمت اصلی است:

- پیدا کردن پارامترهای ثابت، یعنی مختصات دوربین و همچنین شیب زمین.
 - مقداردهی اولیه برای پارامترهایی که تغییر می‌کنند، یعنی ماتریس دوران و فاصله‌ی کانونی.
 - دنبال کردن پارامترهایی که تغییر می‌کنند.
- لازم به ذکر است که در قسمت اول و سوم از الگوریتم دنبال کردن و در قسمت دوم از الگوریتم جستجو استفاده می‌شود.

همچنین در این مقاله روش خوبی برای تشخیص خط بیان شده است که ادعا شده خوب عمل می‌کند.

۲.۱ مدل زمین فوتبال

مدل زمین فوتبال شامل تعدادی خط است. حتی قسمت‌های دایره‌ای خطوط زمین هم به صورت خط صاف در نظر گرفته شده‌اند. به این صورت که هر ۱۰ درجه از دایره‌ها را یک خط صاف در نظر می‌گیریم. توجه کنید که ابعاد زمین فوتبال را به طور دقیق محاسبه کرده و می‌دانیم. به همین خاطر هنگامی که پارامترهای دوربین را داشته باشیم می‌توانیم مدل زمین را بر روی تصاویر دوربین رسم کنیم. از این رسم کردن (انداختن) مدل زمین روی تصاویر در مرحله‌ی جستجو استفاده می‌کنیم.

۲.۲ جستجو

در الگوریتم جستجو ما فرض می‌کنیم که ما پارامترهای ثابت دوربین‌ها را داریم. سپس برای یک تصویر پارامترهایی که تغییر می‌کنند را (ماتریس دوران و فاصله‌ی کانونی) محاسبه می‌کنیم. این محاسبه می‌تواند به خودی خود برای دوربین‌های ثابت مورد استفاده قرار بگیرد یا برای مقدار دهی اولیه در عملیات دنبال کردن دوربین‌های مورد استفاده قرار گیرد. از این به بعد فرض می‌کنیم که فاز جستجو برای مقدار دهی اولیه مورد استفاده قرار می‌گیرد.

برای مقدار دهی اولیه زاویه‌های pan و tilt و مقدار فاصله‌ی کانونی برای یک تصویر خاص، ابتدا فرض می‌کنیم که از تصویر ورودی را تبدیل به تصویر باینری کرده‌ایم که در آن نقاطی که ممکن است روی خطوط قرار داشته باشند در آن مشخص شده است. مرحله‌ی تشخیص خطوط از تبدیل هاف^۳ استفاده می‌کند با این تفاوت که تصویر را به ۱۰ قسمت مساوی (عمودی یا افقی، بسته به نزدیک به عمودی یا نزدیک به افقی بودن خطوط زمین) تقسیم می‌کند و برای هر کدام از آن قسمت‌ها یک تبدیل هاف جدا را محاسبه می‌کند. لازم است که توجه کنید برای هر بار مقدار دهی اولیه، یک بار تبدیل هاف تصویر حساب می‌شود.

سپس برای مقادیر مختلف ممکن از پارامترهای pan، tilt و فاصله‌ی کانونی با اندازه‌ی گام مشخصی، مدل زمین را بر روی تصویر می‌اندازیم. حال برای هر خط از خطوط مدل زمین که در تصویر قابل رؤیت است، بسته به اینکه در کدام نواحی ۱۰ گانه در تقسیم بندی قرار دارد، نقطه‌ی معادل را در فضا‌های هاف ده‌گانه پیدا می‌کنیم و مقدار آن نقاط را در فضای هاف برای تمام خطوط زمین با یکدیگر جمع می‌کنیم. حاصل این جمع به ما ارزش این پارامترهای دوربین را می‌دهد.

در انتهای جستجو از بین تمام پارامترهای مختلف دوربین، آن ۱۰ تایی را انتخاب می‌کنیم که دارای ارزش بیش‌تری هستند. در مراحل دنبال کردن دوربین در حقیقت ۱۰ دنبال کردن همزمان را انجام می‌دهیم. ولی پس از چند مرحله از دنبال کردن مشخص می‌شود که فقط یکی از این مقادیر خوب دنبال می‌شود و آن را انتخاب می‌کنیم. همچنین باید توجه کنیم که در انتخاب ۱۰ تا از سه‌تایی پارامترهای pan، tilt و فاصله‌ی کانونی که دارای بیش‌ترین ارزش هستند، از انتخاب مقادیر نزدیک به هم که عملاً ماکزیمم‌های محلی هستند اجتناب می‌کنیم و آن‌هایی را انتخاب می‌کنیم که دارای تفاوت حداقلی در زوایا باشند.

^۳ Hough transform

۲.۳ دنبال کردن

در مرحله‌ی دنبال کردن، ما یک حدس اولیه از پارامترهایی دوربین که همان مقادیر زمان (فریم) قبل هستند را داریم و حال می‌خواهیم برای فریم جدیدی که دوربین تصویر آن را ضبط کرده است، پارامترهای جدید را بدست آوریم. این مرحله در حقیقت یک روش بهینه‌سازی گام به گام^۴ است که می‌توان آن را برای پارامترهای ثابت، متحرک یا هر دو دسته پارامترهای دوربین حل کرد. برای اطلاع از جزئیات بیشتر این قسمت به خود مقاله مراجعه کنید.

۳ موارد مبهم

این مقاله دو قسمت مبهم برای من دارد:

- اول در مرحله‌ی تعیین شیب زمین و موقعیت دوربین، چگونه مسئله‌ی بهینه‌سازی را حل می‌کند؟
- دوم اینکه در مرحله‌ی جستجو، گام را به چه طریقی تعیین کرده است؟

۴ پیاده‌سازی

به نظر می‌رسد که این مقاله و راه حلی که ارائه داده است، قابلیت حل کردن مسئله‌ی ما را داشته باشد. از آنجایی که در صورت مسئله‌ی ما، موقعیت دوربین‌ها تعیین شده است، در فازهای اول پیاده‌سازی این مقاله نیازی به انجام دادن این کار نیست.

ضمناً اکثر دوربین‌های ما ثابت هستند. این مسئله هم یک حسن است و هم یک عیب. حسن از این نظر که برای واسنجی این دوربین‌ها نیازی به پیاده‌سازی مرحله‌ی دنبال کردن نداریم. و عیب از این نظر که در روش این مقاله خروجی مرحله‌ی جستجو ۱۰ سری پارامتر برای دوربین است که در مرحله‌ی دنبال کردن بهترین آن‌ها تعیین می‌شود که از آنجایی که دوربین ما تکان نمی‌خورد، ما نمی‌توانیم بهترین سری پارامترها را با روش این مقاله انتخاب کنیم. پس باید فکری به حال انتخاب بهترین سری پارامترها بکنیم.

^۴iterative

کتاب نامه

- [1] L. Vacchetti, V. Lepetit, P. Fua: Combining edge and texture information for real-time accurate 3d camera tracking. International Symposium on Mixed and Augmented Reality, pp 48-56, 2004
- [2] T. Watanabe, M. Haseyama, H. Kitajima: A soccer field tracking method with wire frame model from TV images. International Conference on Image Processing, pp 1633-1636, 2004