



دانشگاه زنجان

دانشکده مهندسی - گروه برق

پایان نامه کارشناسی

گرایش الکترونیک

عنوان:

توصیف نقاط کلیدی در تصاویر با تبدیل نا متغیر به تغییر مقیاس

تشکری یلدا

استاد راهنما:

دکتر چرمی

شهریور ۹۴

## چکیده

این مقاله یک روش برای استخراج ویژگی های ثابت متمایز از تصاویر است که می تواند برای انجام تطبیق مطمئن بین دیدگاه های مختلف یک شی یا صحنه مورد استفاده قرار گیرد. که شامل ویژگی های ثابت، مقیاس تصویر و

چرخش هستند، نشان دهنده و ارائه کننده تطبیق مقاوم در سراسر طیف قابل توجهی از اعوجاج، تغییر دیدگاه های

سه بعدی  $3D$ ، علاوه بر این سر و صدا، و تغییردهنده در روشنایی می باشند. ویژگی های بسیار متمایز، به این

معناست که از ویژگی های مجزا می تواند به درستی با احتمال بالا در برابر یک پایگاه داده بزرگ از ویژگی های

بسیاری از تصاویر همسان عمل نماید. این مقاله همچنین یک رویکرد از این ویژگی ها را برای تشخیص شیء

توصیف می نماید.

شناخت ها مطابق با ویژگی های منحصر به فرد یک پایگاه داده با توجه به ویژگی های اشیاء شناخته شده با

استفاده از یک الگوریتم سریع نزدیکترین همسایه، دنبال کردن یک تبدیل هاف جهت شناسایی خوشه متعلق به

یک جسم واحد، و در نهایت انجام راه حل امنیتی از طریق حداقل رسانی مربعات جهت سازگار نمودن پارامترهای

طرح است.

این رویکرد جهت تشخیص استوار می تواند اشیاء درهم و برهم و مسدود را شناسایی نماید در حالی که دستیابی

به عملکرد زمانی واقعی است.  $3D$ ، علاوه بر این سر و صدا، و تغییردهنده در روشنایی می باشند. ویژگی های بسیار متمایز، به این

معناست که از ویژگی های مجزا می تواند به درستی با احتمال بالا در برابر یک پایگاه داده بزرگ از ویژگی های

بسیاری از تصاویر همسان عمل نماید. این مقاله همچنین یک رویکرد از این ویژگی ها را برای تشخیص شیء

توصیف می نماید.

شناخت ها مطابق با ویژگی های منحصر به فرد یک پایگاه داده با توجه به ویژگی های اشیاء شناخته شده با

استفاده از یک الگوریتم سریع نزدیکترین همسایه، دنبال کردن یک تبدیل هاف جهت شناسایی خوشه متعلق به

یک جسم واحد، و در نهایت انجام راه حل امنیتی از طریق حداقل رسانی مربعات جهت سازگار نمودن پارامترهای

طرح است.

این رویکرد جهت تشخیص استوار می تواند اشیاء درهم و برهم و مسدود را شناسایی نماید در حالی که دستیابی

به عملکرد زمانی واقعی است.  $3D$ ، علاوه بر این سر و صدا، و تغییردهنده در روشنایی می باشند. ویژگی های بسیار متمایز، به این

معناست که از ویژگی های مجزا می تواند به درستی با احتمال بالا در برابر یک پایگاه داده بزرگ از ویژگی های

## فهرست مطالب

۱. مقدمه .....	۲
۲- تحقیقات مرتبط .....	۴
۳- تشخیص اکستریم مقیاس فضا .....	۸
۳.۱ تشخیص اکستریم های محلی .....	۱۱
۳.۲ مقیاس فرکانس نمونه برداری .....	۱۳
۳.۳ فرکانس نمونه گیری در حوزه مکان .....	۱۵
۴- محلی سازی نقطه کلیدی دقیق .....	۱۶
۴.۱ حذف پاسخ لبه .....	۱۹
۵- انتساب جهت .....	۲۰
۶- توصیف تصویر محلی .....	۲۳
۶.۱ توصیف نمایندگی .....	۲۵
۶.۲ توصیف تست .....	۲۷
۶.۳ حساسیت به تغییر آفین .....	۲۸
۶.۴ تطبیق پایگاه داده های بزرگ .....	۳۰
۷- برنامه برای تشخیص شی .....	۳۲
۷.۱ تطبیق نکته کلیدی .....	۳۲
۷.۲ شاخص کارآمد نزدیکترین همسایه .....	۳۴
۷.۳ خوشه با تبدیل هاف .....	۳۵
۷.۴ راه حلی برای پارامترهای آفین .....	۳۶
۸- شناخت نمونه (مثال) .....	۴۰
مراجع .....	۴۴

## ۱. مقدمه

تطبيق تصوير جنبه ی اساسی در بسیاری از مشکلات از نظر دیدگاه کامپیوتری، از جمله شی یا تشخیص صحنه، حل ساختار سه بعدی  $3D$  از تصاویر متعدد، مکاتبات استریو، و ردیابی حرکت دارد. این مقاله ویژگی های تصویری را مطرح می کند که دارای خواص بسیاری هستند، که آنها را برای تطبيق متفاوت تصاویری از یک شی یا صحنه مناسب ارائه می نمایند.

ویژگی های ثابت به پوسته پوسته شدن تصویر و چرخش، و تا حدی ثابت به تغییر در نور و نقطه نظر دوربین سه بعدی  $3D$  بستگی دارد. آنها به خوبی در هر دو حوزه مکانی و فرکانس موضعی، آرائه مبنایی برای جسم و تشخیص صحنه شود.

سبب کاهش احتمال اختلال توسط انسداد، درهم و برهمی، و یا سر و صدا. می گردند. تعداد زیادی از ویژگی ها می توان از تصاویر معمولی با الگوریتم های کارآمد استخراج گردد.

علاوه بر این، ویژگی های بسیار متمایز وجود دارد، که اجازه می دهد تا از ویژگی های تکی به درستی با احتمال بالا در برابر یک پایگاه داده بزرگ با ویژگی های همسان استفاده گردد، و سبب آرائه مبنایی برای جسم و تشخیص صحنه شود.

هزینه استخراج این ویژگی ها با در نظر گرفتن یک رویکرد فیلتر آشکار، که در آن عملیات گران تر است فقط در مکان های که آزمون اولیه تصویب گردیده است به حداقل می رسد.

در زیر مراحل عمده محاسبات مورد استفاده برای تولید مجموعه ای از ویژگی های تصویر عبارتند از:

۱. تشخیص اکستریم مقیاس فضا: مرحله اول شامل محاسبات جستجو با بیش از تمام مقیاس ها و مکان های تصویری است. این مسئله کارآمد با استفاده از یک تابع تفاوت گاوسی برای شناسایی نقاط بالقوه ثابت با مقیاس و جهت گیری مناسب اجرا شده است.
۲. محلی سازی نقطه کلیدی: در هر محل کاندید، یک مدل دقیق مناسب برای تعیین محل و مقیاس وجود دارد. نکات کلیدی بر اساس اقدامات پایدار انتخاب شده است.
۳. انتساب جهت: یک یا چند جهت مکانی و نکته کلیدی بر اساس جهت گرادیان به تصویر محلی اختصاص داده شده است. همه عملیات های آینده بر روی داده های تصویر قرار دارد که نسبت به

اختصاص جهت گیری، مقیاس، و محل هر یک از ویژگی تبدیل شده است، در نتیجه سبب ارائه ناپذیری این تحولات شده است.

۴. توصیف نقطه کلیدی: شیب تصویر محلی در مقیاس انتخاب شده در منطقه در اطراف هر نقطه کلیدی اندازه گیری می شود. این مسئله به نمایندگی اجازه می دهد تا برای سطوح قابل توجهی از اعوجاج شکل محلی و تغییر در روشنایی تبدیل شود.

این رویکرد به نام تبدیل ویژگی مقیاس نابسته (غربال کردن)، به عنوان تبدیل داده های تصویر به مختصات مقیاس ثابت نسبت به ویژگیهای محلی است. یک جنبه مهم این روش این است که آن تولید تعداد زیادی از ویژگی ها سبب پر پوشش شدن تصویر در طیف گسترده ای از مقیاس و مکان می گردد. یک تصویر معمولی با اندازه  $500 \times 500$  پیکسل شروع خواهد شد و حدوداً  $2000$

ویژگی های پایدار (اگر چه این تعداد بستگی به هر دو محتوای تصویر و انتخاب برای پارامترهای مختلف) دارد. به خصوص مقداری از ویژگی ها برای تشخیص شیء، که در آن توانایی تشخیص اشیاء کوچک در پس زمینه به هم ریخته است مستلزم آن است که حداقل ۳ ویژگی درست از هر شیء برای شناسایی قابل اعتماد همسان اساسی است. برای تطبیق تصویر و تشخیص آن، غربال کردن ویژگی ها اول از مجموعه ای از تصاویر مرجع استخراج شده و در پایگاه داده ذخیره می گردد. یک

تصویر جدید به صورت جداگانه با مقایسه هر یک از ویژگی تصویر جدید به پایگاه داده قبلی متصل شده و سبب پیدا کردن ویژگی های کاندید و تطبیق آن بر اساس فاصله اقلیدسی از بردارهای ویژگی همسان می گردد. این مقاله الگوریتم های سریع را یافته و نزدیکترین همسایه که می تواند این محاسبات را به سرعت در پایگاه های داده بزرگ انجام دهد را مورد بحث قرار می دهد.

توصیف نکته کلیدی بسیار متمایز هستند، که اجازه می دهد تا از ویژگی های انفرادی برای پیدا کردن قواعد صحیح با احتمال مناسب در یک پایگاه داده بزرگ استفاده می شود.

با این حال، در یک تصویر به هم ریخته، بسیاری از ویژگی های پس زمینه، برای بازی درست در پایگاه داده وجود ندارد، و ظهور بسیاری از قواعد نادرست علاوه بر آنهایی که درست است. قواعد صحیح را می توان از مجموعه ای کامل از بازی ها شناسایی کرد و زیر مجموعه ای از نکات کلیدی

که بر روی جسم و محل آن قرار دارد، مقیاس و جهت گیری در تصویر جدید در همه ی موارد فیلتر



دانشجویان محترم:

جهت دسترسی به متن کامل پایان نامه‌ها به کتابخانه دانشکده مهندسی و یا آزمایشگاه پروژه گروه برق مراجعه فرمایید.

## ۹- نتیجه گیری

نقاط کلیدی SIFT در این مقاله با توجه به تمایز، ما را قادر می سازد اعمال درستی برای یک نقطه کلیدی از یک پایگاه داده بزرگ در رابطه با دیگر نقاط کلیدی انتخاب می شود بویژه مفید است. این تمایز سبب مونتاژ کردن یک بردار با ابعاد بالا به نمایندگی از شیب تصویر در یک منطقه محلی به دست می آید.

نقطه کلیدی نشان دهنده ی ثابت چرخش تصویر و مقیاس قوی در سراسر طیف قابل توجهی از گروه برق آزمایشگاه پروژه برق دانشگاه زنجان علاوه بر این سر و صدا، و تغییر در روشنایی موثر است.

تعداد زیادی از نقاط کلیدی را می توان با تصاویر معمولی، که منجر به قوت در استخراج اشیاء کوچک در میان درهم و برهمی شده است. واقعیت این است که نقاط کلیدی بیش از یک طیف کامل از فلس شناسایی شده اند و بدان معنی است که ویژگی های کوچک و محلی برای تطبیق اشیاء

کوچک و بسیار مسدود در دسترس هستند، در حالی که نقاط کلیدی بزرگ به خوبی برای تصاویر و سر و صدا سبب محاسبه کارآمد است، به طوری که چند هزار نقطه کلیدی را می توان از یک تصویر معمولی با عملکرد نزدیک به زمان واقعی بر روی سخت افزار استاندارد کامپیوترهای شخصی استخراج کرد. در این مقاله نیز روشی برای استفاده از نقاط کلیدی جهت تشخیص شی ارائه شده

است. رویکرد توصیف شده با استفاده از تقریب نزدیکترین همسایه، سبب تبدیل هاف برای شناسایی خوشه است که بر روی جسم در همه موارد مطرح است، تعیین حداقل مربعات، و تایید نهایی را در برخواهد داشت.

دیگر کاربردهای بالقوه شامل مشخصات مطابق با بازسازی ۳D، ردیابی حرکت و تقسیم بندی، محلی سازی ربات، تصویر مونتاژ پانوراما، کالیبراسیون، و هر نوع دیگری است که نیازمند شناسایی و تطبیق

مکان بین تصاویر است. بسیاری از جهات برای تحقیق بیشتر در استخراج ویژگی های تصویر ثابت و مشخص وجود دارد.

تست سیستماتیک بر روی مجموعه داده با دیدگاه و روشنایی تغییرات کامل (سه بعدی) ۳D مورد نیاز است.



## References

Arya, S., and Mount, D.M. ۱۹۹۳. Approximate nearest neighbor queries in fixed dimensions. In

Fourth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA'۹۳), pp. ۲۷۱-۲۸۰.

Arya, S., Mount, D.M., Netanyahu, N.S., Silverman, R., and Wu, A.Y. ۱۹۹۸. An optimal algorithm

for approximate nearest neighbor searching. Journal of the ACM, ۴۵:۸۹۱-۹۲۳.

Ballard, D.H. ۱۹۸۱. Generalizing the Hough transform to detect arbitrary patterns.

Pattern Recognition, ۱۳(۲):۱۱۱-۱۲۲.

Basri, R., and Jacobs, D.W. ۱۹۹۷. Recognition using region correspondences. International Journal

of Computer Vision, ۲۵(۲):۱۴۵-۱۶۶.

Baumberg, A. ۲۰۰۰. Reliable feature matching across widely separated views. In Conference on

Computer Vision and Pattern Recognition, Hilton Head, South Carolina, pp. ۷۷۴-۷۸۱.

Beis, J. and Lowe, D.G. ۱۹۹۷. Shape indexing using approximate nearest-neighbour search in highdimensional

spaces. In Conference on Computer Vision and Pattern Recognition, Puerto Rico, pp. ۱۰۰۰-۱۰۰۶.

Brown, M. and Lowe, D.G. ۲۰۰۲. Invariant features from interest point groups. In British Machine

Vision Conference, Cardiff, Wales, pp. ۶۵۶-۶۶۵.

Carneiro, G., and Jepson, A.D. ۲۰۰۲. Phase-based local features. In European Conference on Computer

Vision (ECCV), Copenhagen, Denmark, pp. ۲۸۲-۲۹۶.

Crowley, J. L. and Parker, A.C. ۱۹۸۴. A representation for shape based on peaks and ridges in the

difference of low-pass transform. IEEE Trans. on Pattern Analysis and Machine Intelligence,

۶(۲):۱۵۶-۱۷۰. ۲۶

Edelman, S., Intrator, N. and Poggio, T. ۱۹۹۷. Complex cells and object recognition.

Unpublished manuscript:

<http://kybele.psych.cornell.edu/~edelman/archive.html>

- Fergus, R., Perona, P., and Zisserman, A. ۲۰۰۳. Object class recognition by unsupervised scaleinvariant learning. In IEEE Conference on Computer Vision and Pattern Recognition, Madison, Wisconsin, pp. ۲۶۴-۲۷۱.
- Friedman, J.H., Bentley, J.L. and Finkel, R.A. ۱۹۷۷. An algorithm for finding best matches in logarithmic expected time. ACM Transactions on Mathematical Software, ۳(۳):۲۰۹-۲۲۶.
- Funt, B.V. and Finlayson, G.D. ۱۹۹۵. Color constant color indexing. IEEE Trans. on Pattern Analysis and Machine Intelligence, ۱۷(۵):۵۲۲-۵۲۹.
- Grimson, E. ۱۹۹۰. Object Recognition by Computer: The Role of Geometric Constraints, The MIT Press: Cambridge, MA.
- Harris, C. ۱۹۹۲. Geometry from visual motion. In Active Vision, A. Blake and A. Yuille (Eds.), MIT Press, pp. ۲۶۳-۲۸۴.
- Harris, C. and Stephens, M. ۱۹۸۸. A combined corner and edge detector. In Fourth Alvey Vision Conference, Manchester, UK, pp. ۱۴۷-۱۵۱.
- Hartley, R. and Zisserman, A. ۲۰۰۰. Multiple view geometry in computer vision, Cambridge University Press: Cambridge, UK.
- Hough, P.V.C. ۱۹۶۲. Method and means for recognizing complex patterns. U.S. Patent ۳,۰۶۹,۶۵۴.
- Koenderink, J.J. ۱۹۸۴. The structure of images. Biological Cybernetics, ۵۰:۳۶۳-۳۹۶.
- Lindeberg, T. ۱۹۹۳. Detecting salient blob-like image structures and their scales with a scale-space primal sketch: a method for focus-of-attention. International Journal of Computer Vision, ۱۱(۳): ۲۸۳-۳۱۸.
- Lindeberg, T. ۱۹۹۴. Scale-space theory: A basic tool for analysing structures at different scales. Journal of Applied Statistics, ۲۱(۲):۲۲۴-۲۷۰.
- Lowe, D.G. ۱۹۹۱. Fitting parameterized three-dimensional models to images. IEEE Trans. on Pattern Analysis and Machine Intelligence, ۱۳(۵):۴۴۱-۴۵۰.
- Lowe, D.G. ۱۹۹۹. Object recognition from local scale-invariant features. In International Conference on Computer Vision, Corfu, Greece, pp. ۱۱۵۰-۱۱۵۷.
- Lowe, D.G. ۲۰۰۱. Local feature view clustering for 3D object recognition. IEEE Conference on

- Computer Vision and Pattern Recognition, Kauai, Hawaii, pp. ۶۸۲-۶۸۸.
- Luong, Q.T., and Faugeras, O.D. ۱۹۹۶. The fundamental matrix: Theory, algorithms, and stability analysis. *International Journal of Computer Vision*, ۱۷(۱):۴۳-۷۶.
- Matas, J., Chum, O., Urban, M., and Pajdla, T. ۲۰۰۲. Robust wide baseline stereo from maximally stable extremal regions. In *British Machine Vision Conference*, Cardiff, Wales, pp. ۳۸۴-۳۹۳.
- Mikolajczyk, K. ۲۰۰۲. Detection of local features invariant to affine transformations, Ph.D. thesis, Institut National Polytechnique de Grenoble, France.
- Mikolajczyk, K., and Schmid, C. ۲۰۰۲. An affine invariant interest point detector. In *European Conference on Computer Vision (ECCV)*, Copenhagen, Denmark, pp. ۱۲۸-۱۴۲.
- Mikolajczyk, K., Zisserman, A., and Schmid, C. ۲۰۰۳. Shape recognition with edge-based features. In *Proceedings of the British Machine Vision Conference*, Norwich, U.K.
- Moravec, H. ۱۹۸۱. Rover visual obstacle avoidance. In *International Joint Conference on Artificial Intelligence*, Vancouver, Canada, pp. ۷۸۵-۷۹۰.
- Nelson, R.C., and Selinger, A. ۱۹۹۸. Large-scale tests of a keyed, appearance-based ۳-D object recognition system. *Vision Research*, ۳۸(۱۵):۲۴۶۹-۸۸.
- Pope, A.R., and Lowe, D.G. ۲۰۰۰. Probabilistic models of appearance for ۳-D object recognition. *International Journal of Computer Vision*, ۴۰(۲):۱۴۹-۱۶۷. ۲۷
- Pritchard, D., and Heidrich, W. ۲۰۰۳. Cloth motion capture. *Computer Graphics Forum (Eurographics)* ۲۰۰۳(۳):۲۶۳-۲۷۱.
- Schaffalitzky, F., and Zisserman, A. ۲۰۰۲. Multi-view matching for unordered image sets, or 'How do I organize my holiday snaps?'" In *European Conference on Computer Vision*, Copenhagen, Denmark, pp. ۴۱۴-۴۳۱.
- Schiele, B., and Crowley, J.L. ۲۰۰۰. Recognition without correspondence using multidimensional receptive field histograms. *International Journal of Computer Vision*, ۳۶(۱):۳۱-۵۰.
- Schmid, C., and Mohr, R. ۱۹۹۷. Local grayvalue invariants for image retrieval. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, ۱۹(۵):۵۳۰-۵۳۴.

- Se, S., Lowe, D.G., and Little, J. ۲۰۰۱. Vision-based mobile robot localization and mapping using scale-invariant features. In International Conference on Robotics and Automation, Seoul, Korea, pp. ۲۰۵۱-۵۸.
- Se, S., Lowe, D.G., and Little, J. ۲۰۰۲. Global localization using distinctive visual features. In International Conference on Intelligent Robots and Systems, IROS ۲۰۰۲, Lausanne, Switzerland, pp. ۲۲۶-۲۳۱.
- Shokoufandeh, A., Marsic, I., and Dickinson, S.J. ۱۹۹۹. View-based object recognition using saliency maps. Image and Vision Computing, ۱۷: ۴۴۵-۴۶۰.
- Torr, P. ۱۹۹۵. Motion Segmentation and Outlier Detection, Ph.D. Thesis, Dept. of Engineering Science, University of Oxford, UK.
- Tuytelaars, T., and Van Gool, L. ۲۰۰۰. Wide baseline stereo based on local, affinely invariant regions. In British Machine Vision Conference, Bristol, UK, pp. ۴۱۲-۴۲۲.
- Weber, M., Welling, M. and Perona, P. ۲۰۰۰. Unsupervised learning of models for recognition. In European Conference on Computer Vision, Dublin, Ireland, pp. ۱۸-۳۲.
- Witkin, A.P. ۱۹۸۳. Scale-space filtering. In International Joint Conference on Artificial Intelligence, Karlsruhe, Germany, pp. ۱۰۱۹-۱۰۲۲.
- Zhang, Z., Deriche, R., Faugeras, O., and Luong, Q.T. ۱۹۹۵. A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry. Artificial Intelligence, ۷۸: ۸۷-۱۱۹.