Формирование панорамных изображений от камер видеорегистрации автомобильного потока

07, июль 2012

DOI: 10.7463/0712.0445483

Тассов К. Л., Федотов А. Л.

УДК.004.93

Россия, МГТУ им. Н.Э. Баумана

tassov@mail.ru

Введение

На сегодняшний день в существующих системах видеорегистрации транспортных средств камеры устанавливаются над полосами дорожного полотна. При частичном попадании автомобиля в поле зрения одной из камер возникает проблема фиксации государственных регистрационных знаков (ГРЗ) автомобилей, проезжающих между полосами, поскольку большинство камер установлено чётко над определённой полосой.

Для решения поставленной проблемы было принято решение формировать панорамное изображение из поступающих изображений с сервера распознавания, располагающихся на одной линии многополосной дороги, которое далее подаётся в модуль распознавания ГРЗ.

Из используемых аппаратно-программных комплексов (АПК) на автомобильных дорогах («Автоураган», «Поток», «Авто-Инспектор») только «Автоураган» предусматривает возможность формирования панорамных изображений. Однако применяемый алгоритм имеет узкий круг применений. Алгоритм заключается в сравнении изображений раскадрованного видеопотока. Если пиксели на смежных изображениях с учётом изменения интенсивности отличаются на 10 %, то совпадение успешно. Данная особенность ограничивает область применения алгоритма только на жёстких установках типа «Ферма» и требует периодического проведения технического обслуживания и настройки программного обеспечения.

Большинство дорог, на которых используется АПК «Автоураган», оборудованы видеокамерами, которые расположены на Г-образных консолях или на растяжках и подвержены к погодным условиям и вибрации, поэтому для формирования панорамного изображения с таких камер необходим другой алгоритм.

1 Алгоритм формирования панорамного изображения

Формирование панорамного изображения состоит из следующих этапов [1]:

- 1) определение особых точек двух соседних изображений;
- 2) нахождение соответствий между особыми точками;
- 3) вычисление матрицы преобразования;
- 4) наложение изображений и получение итогового.

2 Определение особых точек

На первом этапе определяются инвариантные характеристики двух соседних изображений путём выделения особых точек и их дескрипторов. Особая точка m – это точка изображения, окрестность которой $o(m)$ можно отличить от окрестности любой другой точки изображения $o(n)$ в некоторой другой окрестности особой точки $o_2(m)$ [2]. Дескриптор – это идентификатор особой точки, выделяющий её из остального множества особых точек.

Для обнаружения особых точек используется модифицированный алгоритм преобразования масштабно-инвариантных характеристик (англ. *Scale Invariant Feature Transform, SIFT*). Данный алгоритм позволяет определять особые точки в виде капель (англ. *blobs*), так как они инвариантны ко всем преобразованиям (к аффинным преобразованиям, изменениям освещённости, положению камеры и к шуму). Капли – это структуры, описываемые координатами центра, масштабом и направлением. Подобного вида структуры являются самыми сложными, высокоуровневыми среди всех видов форм особых точек, и поэтому обеспечивают устойчивое их обнаружение. Алгоритм *SIFT* показывает лучшие результаты обнаружения особых точек по качеству чёткости и контрастности изображения, чем другие алгоритмы обнаружения капель, и подходит к предметной области [3, 4, 5]. Недостатком *SIFT* является его вычислительная сложность, что ограничивает его применение в режиме постобработки.

Алгоритм *SIFT* разделяется на 5 частей [1, 6]:

- 1) построение пирамиды гауссиан и их разностей. На этом шаге обеспечивается инвариантность к масштабированию;
- 2) определение экстремумов;
- 3) уточнение особых точек;
- 4) определение ориентаций особых точек (обеспечивается инвариантность к повороту);
- 5) построение дескрипторов (обеспечивается инвариантность к освещению, шуму, изменению положения камеры).

На первом шаге алгоритма *SIFT* [6, 7] строится масштабируемое пространство изображения – набор изображений, сглаженных фильтром

Гаусса $G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}$, где (x, y) – координаты точки, σ – радиус размытия. По

ним строится разность гауссиан $D(x, y, \sigma)$ – попиксельное вычитание изображений в одной октаве с разным коэффициентом размытия. Октава – изображение в одном масштабе, размытое фильтром Гаусса (4 изображения в одной октаве). На этом шаге обеспечивается инвариантность к масштабированию. Затем определяются экстремумы, которые заносятся в список потенциальных особых точек.

Далее происходит уточнение особых точек, которое состоит из двух составляющих:

- 1) исключаются точки с малой контрастностью с помощью вычисления экстремума разности гауссиан. Разность гауссиан раскладывается многочленом Тейлора второго порядка, взятого в точке вычисленного экстремума;
- 2) исключаются граничные точки (точки, имеющие большой локальный изгиб вдоль границы и малый в перпендикулярном направлении).

На конечном шаге алгоритма *SIFT* для окрестности особой точки вычисляются изменения яркостей точек, по которым строится дескриптор. Дескриптор – это вектор из 64 чисел, позволяет получить инвариантность относительно положения камеры. Затем дескриптор нормализуется, за счёт чего достигается инвариантность относительно изменения освещения.

3 Нахождение соответствий

На следующем этапе формирования панорамного изображения между особыми точками находятся соответствия на основе полученных инвариантных дескрипторов. Для этого строится kd -дерево – это структура, предназначенная для хранения конечного множества точек k -мерного пространства [8]. В результате выделяется набор пар взаимосвязанных особых точек.

4 Вычисление матрицы преобразования

Полученная последовательность может иметь ложные соответствия. Для удаления ложных соответствий применяется метод согласования случайных выборок *RANSAC* (англ. *Random Sample Consensus*) – итерационный, вероятностный метод, в котором определяется минимальная симметрическая погрешность между парами точек, после чего вычисляются коэффициенты матрицы перспективного преобразования H размерности 3×3 (матрица гомографии) методом прямого линейного преобразования *DLT* (англ. *Direct Linear Transformation*). Решается система линейных алгебраических уравнений, в результате определяются 8 коэффициентов (параметров *DLT*), показывающих связь между системами координат плоскостей двух изображений [9].

5 Наложение изображений

Заключительный этап наложения изображений состоит из следующих частей:

- 1) определяется размер итогового панорамного изображения;
- 2) первое изображение без изменения копируется в плоскость итогового;
- 3) второе изображение накладывается на плоскость первого с помощью полученной матрицы преобразования;
- 4) общая часть накладывается путём линейной интерполяции.

6 Результаты

Для проверки качества работы алгоритма было проведено исследование влияния размера области перекрытия изображений на правильность формирования и выяснено, что область перекрытия камер по горизонтали возможно снизить с первоначальных 15 % до 10 %. Размер выборки составляет 5500 изображений.

Заключение

В представленной работе был произведён анализ предметной области, алгоритмов выделения особых точек, проанализированы и реализованы алгоритмы, необходимые для формирования панорамного изображения. Ограничением разработанного программного продукта является обработка не более двух изображений (при большем числе изображений проявляется эффект дисторсии).

К основным направлениям развития следует отнести:

- модификацию алгоритмов для использования в многопроцессорных системах;
- формирование более двух изображений в панорамное.

Пример результата формирования панорамного изображения приведён на следующем рисунке.

Список литературы

- 1) Brown M., Lowe D.G. Automatic Panoramic Image Stitching using Invariant Features // International Journal of Computer Vision. 2007. Vol. 74. No. 1. PP. 59-73. DOI: 10.1007/s11263-006-0002-3;

- 2) Конушин А.С. Слежение за точечными особенностями сцены (Point feature tracking) // Компьютерная графика и мультимедиа. Электрон. журн. 2003. №1(5). Режим доступа <http://cgm.computergraphics.ru/content/view/54> (дата обращения 07.06.2012);
- 3) Juan L., Gwun O. A Comparison of SIFT, PCA-SIFT and SURF // International Journal of Image Processing (IJIP). 2009. Vol. 3. No. 4. PP. 143-152;
- 4) Tran P. Pyramid Matching Using SURF and SIFT Descriptors for SVM Classification. University of Kansas. – 2010. Режим доступа: http://www.ittc.ku.edu/~potetz/EECS_841_Fall09/Lectures/Final%20Projects/Paper_04.pdf (дата обращения 12.06.2012);
- 5) Valgren C., Lilienthal A.J. SIFT, SURF & Seasons: Appearance-based Long-term Localization in Outdoor Environments // Robotics and Autonomous Systems. 2010. Vol. 58. No. 2. PP. 149–156. DOI: <http://dx.doi.org/10.1016/j.robot.2009.09.010>;
- 6) Meng Y., Tiddeman B. Implementing the Scale Invariant Feature Transform (SIFT) Method. Citeseer. – 2008. Режим доступа: http://www.cs.st-andrews.ac.uk/~yumeng/yumeng-SIFTreport-5.18_bpt.pdf (дата обращения 10.06.2012);
- 7) Гонсалес Р., Вудс Р. Цифровая обработка изображений. Пер. с англ. под ред. П.А. Чочиа. М.: Техносфера, 2005. 1072 с.;
- 8) Kybic J., Vnucko I. Approximate Best Bin First k-d Tree All Nearest Neighbor Search with Incremental Updates. Czech Technical University in Prague. – 2010. Режим доступа: <ftp://cmp.felk.cvut.cz/pub/cmp/articles/kybic/Kybic-TR-2010-10.pdf> (дата обращения 10.06.2012);
- 9) Hartley R.I., Zisserman A. Multiple View Geometry in Computer Vision. 2nd ed. Cambridge: Cambridge University Press, 2003. 672 p.

Formation of panoramic images from videocameras recording automobile stream

07, July 2012

DOI: 10.7463/0712.0445483

Tasov K.L., Fedotov A.L.

Russia, Bauman Moscow State Technical University

tasov@mail.ru

The article considers creating panoramic images in the data coming from the camera video recording automobile stream. The authors describe each of the steps and give a description of the sequence of actions to achieve the final result. The analysis of the results is also provided.

Publications with keywords: [kd-tree](#), [point feature \(keypoint\)](#), [panoramic image](#), [perspective transformation](#), [descriptor](#)

Publications with words: [kd-tree](#), [point feature \(keypoint\)](#), [panoramic image](#), [perspective transformation](#), [descriptor](#)

References

1. Brown M., Lowe D.G. Automatic panoramic image stitching using invariant features. *International Journal of Computer Vision*, 2007, vol. 74, no. 1, pp. 59-73. DOI: 10.1007/s11263-006-0002-3.
2. Konushin A.S. Slezhenie za tochechnymi osobennostiami stsny [Point feature tracking]. *Komp'yuternaia grafika i mul'timedia* [Computer Graphics and Multimedia], 2003, no. 1 (5). Available at: <http://cgm.computergraphics.ru/content/view/54>. Accessed June 07, 2012.
3. Juan L., Gwun O. A comparison of SIFT, PCA-SIFT and SURF. *International Journal of Image Processing (IJIP)*, 2009, vol. 3, no. 4, pp. 143-152.
4. Tran P. *Pyramid matching using SURF and SIFT descriptors for SVM classification*. University of Kansas, 2010. Available at: http://www.ittc.ku.edu/~potetz/EECS_841_Fall09/Lectures/Final%20Projects/Paper_04.pdf. Accessed June 12, 2010.
5. Valgren C., Lilienthal A.J. SIFT, SURF & seasons: Appearance-based long-term localization in outdoor environments. *Robotics and Autonomous Systems*, 2010, vol. 58, no. 2, pp. 149-156. DOI: <http://dx.doi.org/10.1016/j.robot.2009.09.010>.

6. Meng Y., Tiddeman B. *Implementing the scale invariant feature transform (SIFT) method*. Citeseer, 2008. Available at: http://www.cs.st-andrews.ac.uk/~yumeng/yumeng-SIFTreport-5.18_bpt.pdf. Accessed June 10, 2012.
7. Gonzalez R.C, Woods R.E. *Digital image processing using MATLAB*. Englewood Cliffs, Pearson Prentice-Hall, 2003. 609 p. (Russ. ed.: Gonsales R., Vuds R. *Tsifrovaia obrabotka izobrazhenii*. Moscow, Tekhnosfera Publ., 2005. 1072 p.).
8. Kybic J., Vnucko I. *Approximate best bin first k-d tree all nearest neighbor search with incremental updates*. Czech Technical University in Prague, 2010. Available at: <ftp://cmp.felk.cvut.cz/pub/cmp/articles/kybic/Kybic-TR-2010-10.pdf>. Accessed June 10, 2012.
9. Hartley R.I., Zisserman A. *Multiple view geometry in computer vision*. 2nd ed. Cambridge, CUP Publ., 2003. 672 p.