

ИКОНИКА – НАУКА ОБ ИЗОБРАЖЕНИИ

УДК 004.93'12, 004.93'14

СОПОСТАВЛЕНИЕ ИЗОБРАЖЕНИЙ ТРЕХМЕРНЫХ СЦЕН С ПОМОЩЬЮ КЛАСТЕРИЗАЦИИ СОПОСТАВЛЕННЫХ ЛОКАЛЬНЫХ ПРИЗНАКОВ ПОСРЕДСТВОМ ПРЕОБРАЗОВАНИЯ ХАФА

© 2014 г. Р. О. Малашин, аспирант

Государственный оптический институт им. С.И. Вавилова, Национальный исследовательский университет информационных технологий, механики и оптики, Санкт-Петербург

E-mail: malashinroman@mail.ru

Приведено описание алгоритмов сопоставления изображений произвольных трехмерных сцен с помощью кластеризации сопоставленных ключевых точек посредством преобразования Хафа. Метод основан на известном методе обнаружения объектов, но предлагается альтернативный способ верификации кластеров сопоставленных ключевых точек. Приведены результаты экспериментов для разных типов ключевых точек, подтверждающие значительное преимущество предлагаемого метода по сравнению с использованием фундаментальной матрицы.

Ключевые слова: сопоставление трехмерных сцен, локальные признаки, преобразование Хафа.

Коды OCIS: 100.3008, 100.5760.

Поступила в редакцию 05.03.2014.

Введение

Автоматическое распознавание трехмерных сцен может быть чрезвычайно полезно в таких задачах как навигация низколетящего беспилотного летательного аппарата (БПЛА) и наземного робота. В данной работе рассматривается способ распознавания изображений трехмерных сцен с помощью локальных признаков.

Такие методы, как правило, включают следующие три этапа обработки изображения:

1. Поиск ключевых точек на изображениях.
2. Построение дескрипторов (описаний) на участках изображений в области ключевых точек.
3. Сопоставление ключевых точек двух изображений.

В настоящее время существует большое количество алгоритмов, предназначенных для обнаружения и описания ключевых точек, которые обладают достаточно хорошими характеристиками с точки зрения робастности

к изменчивости изображений. Тем не менее, при анализе трехмерных сцен большие трудности могут создавать окклюзии, которые приводят к тому, что большинство обнаруженных на первом изображении ключевых точек не могут иметь аналога на втором изображении. Дополнительные сложности накладывают сезонно-суточные изменения и дисторсии. Все это приводит к тому, что даже при использовании наиболее робастных алгоритмов обнаружения и описания ключевых точек подавляющая часть сопоставлений, полученных путем сравнения дескрипторов в пространстве признаков, все равно оказываются ошибочными. Итоговый результат в большой степени зависит от надежности алгоритма удаления ошибок, поэтому основные усилия в ходе исследований были сконцентрированы на этом этапе обработки изображения.

Известен способ удаления ошибочных сопоставлений с помощью восстановления эпиполярной геометрии. Эксперименты показывают,

что такой подход применим только, если в исходном наборе сопоставлений не меньше 25–30% правильных сопоставлений, в то время как для достаточно “сложных” пар требуется устойчивая работа при гораздо большем количестве ошибок.

В литературе также достаточно часто встречается подход, когда взаимное положение всех сопоставленных точек описывается набором локальных преобразований. Для этого каким-либо образом выделяют кластеры элементов, имеющие схожие параметры.

В [1] ключевые точки, обнаруженные на изображениях, кластеризуются с помощью алгоритма ISODATA по положению. Таким образом, осуществляется попытка выделить группы ключевых точек принадлежащие отдельным предметам. Для схожих целей в [2] авторы используют текстурный анализ. После кластеризации сопоставление ключевых точек (структурных элементов) можно производить внутри отдельных подобластей, а окончательное решение принимать по количеству сопоставленных регионов изображений.

В [3] предлагается способ получения кластеров ключевых точек, основанный на принципе минимальной длины описания [4] – параметры локальных преобразований подбираются так, чтобы описание положения ключевых точек на двух изображениях занимало как можно меньшее количество бит информации.

Существует эффективный способ обнаружения плоских объектов известной формы с помощью кластеризации контурных элементов посредством преобразования Хафа [5]. Основная идея этого метода заключается в том, что непрерывное пространство возможных положений объекта представляется в виде дискретного множества (так называемого аккумулятора Хафа), а каждый элемент контурного изображения голосует за определенное положение искомого объекта в пространстве возможных положений. Голоса контурных элементов, принадлежащих искомому образу, аккумулируются в определенной ячейке, а голоса других контуров “размываются” в выбранном пространстве и распределяются равномерно по всем ячейкам. Одним из важных качеств преобразования Хафа является низкая вычислительная сложность по сравнению с другими методами кластеризации.

Если речь идет о ключевых точках, то можно кластеризовать не отдельные точки, а сопо-

ставления пар ключевых точек (сделанные на основе дескрипторов), что существенно уменьшает количество голосов в аккумулятор Хафа, а также позволяет обнаруживать неплоские объекты [6]. Одним из видимых преимуществ данного подхода является то, что учитывается ориентация ключевой точки и размер участка, на котором построен дескриптор, в то время как большинство других методов эту информацию игнорируют и рассматривают сопоставляемые элементы как безразмерные точки в пространстве изображения. В данной работе показывается, что такой подход применим при распознавании произвольных трехмерных сцен, и предлагается альтернативный способ верификации кластеров сопоставленных ключевых точек, позволяющий добиться лучших результатов при формировании уточненной гипотезы, чем метод, основанный на решении системы линейных алгебраических уравнений.

Разработанные алгоритмы способны устойчиво работать, даже если 95% дескрипторов сопоставлены неправильно.

Предварительная обработка сопоставлений

Соотношение правильных и неправильных сопоставлений является ключевым фактором, влияющим на вероятность успешного применения структурного анализа, поэтому перед кластеризацией целесообразно провести предварительную обработку сопоставлений с целью уменьшения количества ошибок. Для проведения исследования свойств различных методов на нескольких изображениях, сопоставленных методом Speeded Up Robust Features (SURF) [7], вручную были отмечены правильные сопоставления. Всего было обработано 18 пар изображений “сложных” трехмерных сцен (более 24000 сопоставленных ключевых точек). Правильными считались сопоставления, дескрипторы которых были посчитаны на участках изображений соответствующих друг другу объектов.

Одним из известных методов удаления ошибочных сопоставлений является использование соотношения расстояний до двух ближайших соседей в многомерном пространстве признаков в качестве показателя уникальности найденного сопоставления. Чем больше соотношение расстояний от одного из дескрипторов первого изображения до двух ближайших к нему

дескрипторов второго изображения, тем вероятность того, что сопоставление этого дескриптора верное, выше [6] (метод 1). Распределения правильных и неправильных сопоставлений в зависимости от выбранного порога представлены на рис. 1.

Исследования показывают, что задание более жесткого порога обеспечивает лучшее соотношение правильных и неправильных сопоставлений, но, как видно из графика, приводит к потере большей части правильных сопоставлений. Оказалось, что лучшие результаты по сравнению с использованием этого метода могут быть получены при удалении сопоставлений, которые ссылаются на одну и ту же ключевую точку второго изображения, за исключением сопоставления с наименьшим расстоянием между дескрипторами в пространстве признаков (метод 2). Удаленные таким образом сопоставления, во-первых, с большей вероятностью, чем другие сопоставления являются

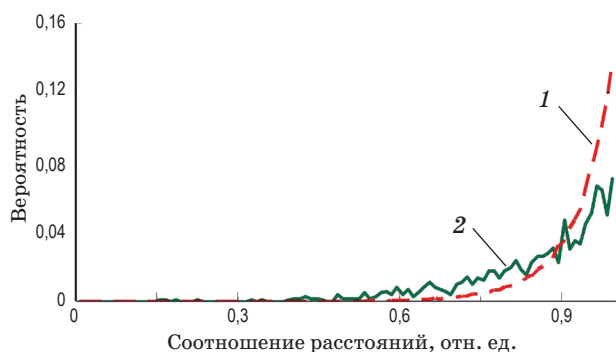


Рис. 1. Распределение вероятности правильных и ошибочных сопоставлений в зависимости от отношения расстояний до ближайших соседей. Кривая 1 – распределение ошибочных сопоставлений, кривая 2 – распределение правильных сопоставлений.

ошибочными, а, во-вторых, с большей долей вероятности, чем другие ошибки, образуют ложный кластер с релевантными взаимными параметрами ключевых точек. График, изображенный на рис. 2, подтверждает это – количество верифицированных ложных кластеров (из расчета на одно сопоставление) значительно снижается. Алгоритмы кластеризации и верификации кластерных гипотез приведены в следующих главах.

В табл. 1 приведены результаты сравнения двух методов с точки зрения уменьшения количества ошибок.

Порог соотношения ближайших соседей 0,77 соответствует лучшему значению F1Score для размеченных сопоставлений. Результаты для порога 0,91 приведены для сравнения метода 1 и метода 2. В этом случае значения Precision и Recall для двух методов приблизительно равны, что говорит о схожих характеристиках с точки зрения уменьшения количества

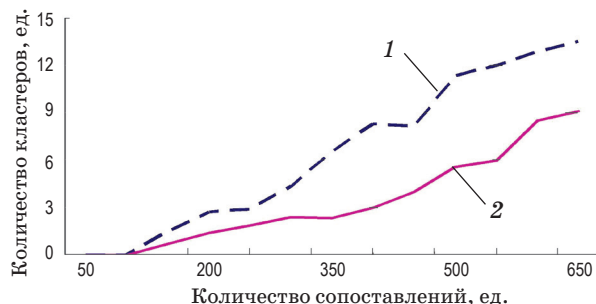


Рис. 2. Количество ложных кластеров из трех и более сопоставлений, прошедших верификацию. Кривая 1 определяет результаты, полученные по сопоставлениям без обработки, кривая 2 – результаты после удаления сопоставлений “многие-к-одному”.

Таблица 1. Предварительная обработка сопоставлений.

№	Метод	tp	fp	tn	fn	Precision	Recall	F1Score
1	б/о	961	23295	0	0	0,039	1	0,076
2	M1(0,77)	231	1202	22093	730	0,161	0,240	0,192
3	M1(0,91)	595	8577	14718	366	0,064	0,619	0,117
4	M2	584	8003	15292	377	0,068	0,607	0,122
5	M2 + M1(0,82)	241	1089	22206	720	0,181	0,250	0,210

Примечание.

В табл. 1 использованы следующие обозначения: б/о – без обработки, M1(C) – после использования соотношения ближайших соседей, где C – использованный порог, M2 – после удаления сопоставлений многие-к-одному, tp – правильные сопоставления (true positives), fp – ошибки первого рода (false positives), tn – неправильные сопоставления (true negatives), fn – ошибки второго рода (false negatives),

Precision = $tp / (tp + fp)$,

Recall = $tp / (tp + fn)$,

F1Score = $2 \times (Precision \times Recall) / (Precision + Recall)$.

ошибок. Таким образом, предпочтительней использовать метод 2, поскольку такой подход дополнительно позволяет сократить количество ложных кластеров. Косвенно о пользе использования метода 2 говорит также и то, что его сочетание с методом 1 позволяет увеличить показатель F1Score, однако в данной работе метод 1 не использовался вообще, чтобы не обеднять результаты сопоставлений.

Кластеризация сопоставленных ключевых точек с помощью преобразования Хафа

Для кластеризации сопоставлений используется алгоритм [6], далее приводится краткое его описание. Д. Лове рассматривает задачу обнаружения объектов на произвольном изображении. В этом случае в качестве модели выступает изображение объекта на темном фоне, требуется определить положение модели на тестовом изображении.

Сопоставление дескрипторов в пространстве признаков дает исходный набор сопоставлений. Как уже было сказано, в нем содержится большое количество ошибок, некоторое количество которых удастся удалить за счет проведения предварительной обработки, описанной выше. Удаление оставшихся ошибок достигается за счет кластеризации локальных преобразований группы подобия, задаваемых сопоставленными ключевыми точками. Такая кластеризация возможна, поскольку каждая из сопоставленных ключевых точек имеет положение, размер и ориентацию, а параметры взаимного масштаба, разворота и смещения задают уникальное преобразование группы подобия. Кластеризация осуществляется путем проведения преобразования Хафа – каждое сопоставление голосует в определенную ячейку аккумулятора Хафа. Аккумулятор имеет четыре измерения, соответствующих углу разворота модели, изменению масштаба и положению модели на плоскости (по осям X и Y). Поскольку изображение объекта может быть подвержено более сложным преобразованиям (например, группе аффинных преобразований) чем преобразование группы подобия, то допускаются значительные отклонения от него. Это достигается за счет использования больших ячеек аккумулятора: 30 градусов по углу поворота, 0,25 максимального разрешения для положения модели и степень 2 для масштаба. Чтобы избежать гранич-

ного эффекта каждое из сопоставлений голосует во все смежные ячейки аккумулятора Хафа.

После этого рассматриваются только те сопоставления, которые попали в ячейки с достаточным количеством голосов и, таким образом, большая часть ошибок удаляется после этого шага. В данной работе рассматривались ячейки, в которые попало не меньше 1% от общего количества сопоставлений (но не меньше четырех).

Сопоставление произвольных трехмерных сцен можно приравнять к сопоставлению многих объектов модельного изображения со многими объектами тестового изображения. Разные объекты могут давать несколько пиков в аккумуляторе Хафа.

Выбор размера ячеек аккумулятора в большей степени зависит не от того, насколько сильные отклонения от преобразования группы подобия допустимы, а от количества сопоставлений при кластеризации, поскольку аффинное преобразование может быть аппроксимировано как одним преобразованием группы подобия с большими отклонениями, так и несколькими преобразованиями группы подобия с небольшими отклонениями. В данном исследовании использовались ячейки вдвое меньшие, чем в работе [6].

Поскольку одним из измерений в пространстве Хафа является положение модели, то необходимо пояснить, что является положением модели в случае трехмерных сцен. Поскольку при кластеризации важно лишь то, что сопоставления соответствующие одному и тому же объекту должны попасть в одну и ту же ячейку аккумулятора, а конкретное положение не важно, то можно предположить, что каждый объект занимает все изображение и его центру соответствует центр изображения. В данной работе для предсказания положения модели $[X', Y']$ по паре сопоставленных ключевых точек p_1 и p_2 использовались следующие формулы:

$$X' = S \cos \theta \cdot dx - S \sin \theta \cdot dy + x_2, \quad (1)$$

$$Y' = S \sin \theta \cdot dx + S \cos \theta \cdot dy + y_2, \quad (2)$$

где

$$dx = X - x_1,$$

$$dy = Y - y_1,$$

$$\theta = \alpha_2 - \alpha_1,$$

$$S = S_2 / S_1,$$

где в свою очередь $[X, Y]$ – центр модельного изображения, $[x_1, y_1]$, α_1 , S_1 – координаты, ориентация и размер ключевой точки p_1 ,

обнаруженной на модельном изображении, $[x_2, y_2]$, α_2 , S_2 – координаты, ориентация и размер ключевой точки p_2 , обнаруженной на тестовом изображении.

Верификация кластерной гипотезы с помощью метода наименьших квадратов

После заполнения аккумулятора, как правило, образуется до нескольких сот ячеек с количеством сопоставлений больше четырех (при сопоставлении изображений с разрешением 640×480). Во многом это объясняется “дублированием” кластеров за счет голосования в смежные ячейки аккумулятора, но, тем не менее, количество ложных кластеров достаточно велико, и необходима дополнительная процедура верификации кластерной гипотезы. В [6] верификация сопоставлений в пиках аккумулятора осуществляется за счет нахождения параметров аффинного преобразования по координатам сопоставленных ключевых точек с помощью метода наименьших квадратов (МНК).

Восстановив параметры вращения, сдвига, масштаба аффинного преобразования, можно отбросить еще некоторое количество ошибочных сопоставлений из кластера, проверяя их на более жесткие ограничения, чем использовались при кластеризации. Процедура повторяется итеративно, что позволяет удалять ошибки постепенно и находить более точные решения, но значительно увеличивает время верификации.

В ходе исследований было выявлено, что при сопоставлении произвольных трехмерных поверхностей (в отличие от плоскостей) лучше подходит ограничение, задаваемое преобразованием группы подобия, а не более общим аффинным преобразованием. Это можно объяснить тем, что по сравнению с аффинным преобразованием нахождения требуют лишь четыре неизвестных, а не шесть. Эти выводы аналогичны тем, что были получены в [8] при решении задачи сопоставления трехмерных объектов. Система линейных алгебраических уравнений (СЛАУ) для преобразования группы подобия имеет вид

$$\begin{bmatrix} x_i & -y_i & 1 & 0 \\ y_i & x_i & 0 & 1 \\ & & \dots & \end{bmatrix} \begin{bmatrix} m \\ n \\ t_x \\ t_y \end{bmatrix} = \begin{bmatrix} u_i \\ v_i \\ \cdot \\ \cdot \\ \cdot \end{bmatrix}, \quad (3)$$

где $[u_i, v_i]$ и $[x_i, y_i]$ – это пары сопоставленных точек, $m = S \cos \theta$, $n = S \sin \theta$, t_x, t_y – параметры преобразования группы подобия.

Тем не менее, данный подход не работает при количестве выбросов в наборе близком к 50% и в некоторых случаях даже при меньшем их количестве, т. к. целевая функция имеет вид:

$$F_{\text{obj}}(m, n, t_x, t_y) = \min \sum_i \left[(x'_i - u_i)^2 + (y'_i - v_i)^2 \right],$$

где $[x'_i, y'_i]$ – координаты точки $[x_i, y_i]$, полученные с помощью параметров преобразования группы подобия m, n, t_x, t_y . Таким образом, выброс со значительным отклонением будет давать намного более значительный вклад в решение, чем сумма отклонений нескольких правильных сопоставлений (пропорционально квадрату отклонения), поэтому решение будет иметь значительные искажения, особенно если в наборе немного сопоставлений.

Некоторого увеличения устойчивости к выбросам можно добиться за счет более детальной проработки системы штрафов за отклонение от предсказываемого положения. Глядя на формулы (1) и (2), нетрудно заметить, что, если ключевая точка будет обнаружена в центре изображения модели, то d_x и d_y будут стремиться к нулю, и предсказываемое положение центра модели будет зависеть только от положения точек. Таким образом, вероятность появления ложного кластера при наличии сопоставлений в центре изображения сильно возрастает, поскольку предсказываемое положение модели не будет зависеть от ошибок в определении взаимного масштаба и взаимного разворота ключевых точек. Поскольку пики аккумулятора Хафа анализируются дополнительно, ложные кластеры можно удалить, но, тем не менее, необходимо учитывать этот эффект при верификации кластерной гипотезы.

Во-первых, вместо того чтобы задавать жесткий порог T на отклонение положения модели, предсказываемого с помощью пары сопоставленных ключевых точек, от положения модели, предсказываемого с помощью параметров преобразования группы подобия, предлагается использовать порог, связанный с расстоянием между точкой модели (центр изображения при кластеризации) и положением ключевой точки

$$T = k \text{Dist} + \lambda, \quad (4)$$

где Dist – расстояние между точкой, обнаруженной на модельном изображении, и центром

изображения (модельной точкой), k – коэффициент допустимого отклонения, а λ – минимальный порог. В данной работе использовались $k = 0,3$, λ – одна сотая от диагонали тестового изображения.

Во-вторых, преобразование дает более точное предсказание о положении точки в тех областях изображения, в которых находились ключевые точки, по которым были рассчитаны параметры, а ключевая точка предсказывает положение модели тем точнее, чем ближе к ней находится модельная точка. В качестве модельной точки предлагается выбирать точку $[X_M, Y_M]$

$$X_M = (X_{cl} + x_1) / 2,$$

$$Y_M = (Y_{cl} + y_1) / 2,$$

где $[x_1, y_1]$ – положение ключевой точки на модельном изображении, а $[X_{cl}, Y_{cl}]$ – координаты кластера, которые можно вычислить по следующей формуле

$$X_{cl} = \frac{1}{N} \sum_{i=1}^N x_i, \quad Y_{cl} = \frac{1}{N} \sum_{i=1}^N y_i,$$

где $[x_i, y_i]$ – координаты N ключевых точек на модельном изображении, входящие в рассматриваемый кластер сопоставлений.

Верификация кластерной гипотезы с помощью итеративных стохастических методов

В ходе исследований были проведены попытки добиться увеличения робастности процедуры верификации кластерной гипотезы за счет использования следующей информации о локальных преобразованиях группы подобия, задаваемых сопоставленными парами ключевых точек:

1. В СЛАУ (3) можно уменьшить количество неизвестных до трех, если получить значение угла разворота как медианный взаимный разворот всех сопоставленных точек, попавших в кластер.

2. Преобразование группы подобия для всего кластера можно получить как медианное локальное преобразование, задаваемое каждым отдельным сопоставлением.

Два указанных выше метода с одной стороны помогают снизить влияние выбросов, но с другой стороны в меньшей степени, чем исходный метод учитывают взаимное положение сопоставленных ключевых точек при расчете

параметров преобразования, что в конечном итоге не позволяет добиться желаемых показателей. Гораздо лучшие результаты с точки зрения надежности и аккуратности восстановленных параметров преобразования могут дать итеративные стохастические методы.

Итеративные стохастические методы (RANSAC [9]) основаны на построении гипотезы по случайной выборке из совокупности исходных данных и выполняются итеративно в два этапа. На первом этапе осуществляется построение гипотезы по нескольким сопоставлениям, случайно выбранным из всей совокупности сопоставлений. На втором этапе осуществляется проверка гипотезы – оценивается, какое количество оставшихся сопоставлений из всей совокупности удовлетворяет выбранной гипотезе. Если выбранные сопоставления окажутся правильными, то получатся адекватные значения параметров и большее количество сопоставлений должно согласиться с этими параметрами. Главным недостатком такого метода является вычислительная сложность, поскольку процедуру надо повторять многократно, а проверка гипотезы может занимать много времени.

Главными параметрами, влияющими на количество необходимых итераций, является соотношение правильных и ошибочных сопоставлений в исходной совокупности данных и количество сопоставлений, выбираемых случайно для расчета параметров. Параметры преобразования группы подобия – масштаб S , вращение θ , смещение $[t_x, t_y]$ – можно однозначно восстановить по двум сопоставлениям, используя следующие формулы:

$$S = \sqrt{\frac{(x_{21} - x_{22})^2 + (y_{21} - y_{22})^2}{(x_{11} - x_{12})^2 + (y_{11} - y_{12})^2}},$$

$$\theta = \arctg\left(\frac{x_{21} - x_{22}}{y_{21} - y_{22}}\right) - \arctg\left(\frac{x_{11} - x_{12}}{y_{11} - y_{12}}\right),$$

$$t_x = x_{22} - (Sx_{21} \cos\theta - Sy_{21} \sin\theta),$$

$$t_y = y_{22} - (Sx_{21} \sin\theta + Sy_{21} \cos\theta),$$

где $[x_{11}, y_{11}]$, $[x_{21}, y_{21}]$ и $[x_{12}, y_{12}]$, $[x_{22}, y_{22}]$ – координаты первой и второй пары сопоставленных ключевых точек, соответственно.

Для увеличения быстродействия процедуры верификации кластерной гипотезы были приняты следующие шаги:

1. После восстановления параметров преобразования группы подобия по двум случайно

выбранным сопоставлениям осуществлялась проверка на согласованность полученного преобразования с преобразованиями, задаваемыми каждым из использованных сопоставлений. Для этого сравнивались восстановленные параметры разворота и масштаба с взаимным разворотом и взаимным масштабом сопоставленных ключевых точек, которые были использованы при расчете параметров. Кроме того, проверялось и то, что указанные параметры двух сопоставлений согласуются между собой. Если согласованности по любому из указанных аспектов нет, то это говорит о том, что одно или оба выбранных сопоставления ошибочны, и поэтому выбиралась другая пара сопоставлений. В этом случае подсчет количества сопоставлений, поддерживающих гипотезу, не проводился, и за счет этого была достигнута значительная экономия вычислительных ресурсов. Использовались следующие пороги: допустимое отклонение по ориентации – 15 градусов, коэффициент отклонения предсказываемого положения модели $k = 0,3$ (см. формулу (4)), отклонение по масштабу $\sqrt{2}$.

2. Поскольку процедура RANSAC в отличие от МНК не чувствительна к отдельным выбросам (если их не слишком много), то нет необходимости решать задачу восстановления параметров итеративно, постепенно удаляя ошибки из кластера, как это предлагается в [6]. Сразу были установлены жесткие ограничения на допустимые отклонения локальных преобразований, задаваемых отдельными сопоставлениями, попавшими в кластер, от параметров преобразования, подобранных RANSAC. Наилучших результатов удалось добиться при выполнении следующей последовательности действий. Сначала выполняется процедура RANSAC (удаляются сопоставления, положение ключевых точек которых от предсказываемого больше 0,1 от диагонали тестового изображения). При этом, если количество сопоставлений, поддерживающих найденные параметры, оказывается меньше 30%, то весь кластер считается ошибочным. Поскольку, как правило, подавляющая часть кластеров ложные, то это позволило сократить количество более сложных проверок. По оставшимся сопоставлениям составлялась СЛАУ (3). Параметры преобразования группы подобия, полученные в результате ее решения с помощью МНК, использовались для проведения проверок сопоставлений по ориентации (отклонения до 15 градусов), масштабу ($\sqrt{2}$), смещению ($k = 0,3$).

3. Если сопоставляемые сцены достаточно просты, то “правильные” кластеры могут оказаться достаточно большими (до нескольких сот сопоставлений). Этих кластеров также оказывается несколько десятков и их содержимое часто практически идентично (из-за голосования сопоставлений во все смежные ячейки), что ведет к тому, что решение по сути одинаковых СЛАУ, согласно п. 2, может занимать много времени. Для того чтобы сэкономить время, перед проведением верификации все кластеры сортируются по размеру. Верификация осуществляется последовательно для каждого кластера, а все сопоставления прошедшие верификацию из последующих кластеров удаляются. Такой подход помимо выигрыша во времени обеспечивает также обогащение результатов сопоставления, поскольку существует вероятность найти внутри кластера второй менее значительный кластер с другими параметрами. Такая ситуация возможна, когда, например, на втором снимке один объект сцены был незначительно сдвинут относительно другого в пространстве. В этом случае при кластеризации сопоставления ключевых точек, обнаруженных на этих объектах, попадут в одни и те же ячейки. Объект, которому соответствует большее количество сопоставлений, всегда будет “подавлять” меньший при верификации кластера, если предварительно не удалить сопоставления, соответствующие большему объекту.

Анализ сложности алгоритма

Для того чтобы в результате выполнения стохастической итеративной процедуры получить правильно решение, необходимо, чтобы хоть одна пара выбранных случайно сопоставлений оказалось правильной. Вероятность α того, что в результате получения k случайных наборов по p сопоставлений ни один набор не окажется правильным

$$\left[1 - (f)^p\right]^k < \alpha,$$

где f – вероятность того, что случайно выбранное сопоставление окажется правильным

$$f = m / n,$$

где m – количество правильных сопоставлений, а n – общее число сопоставлений. В нашем случае для получения прироста в качестве работы алгоритма требуется устойчивое срабатывание при 50% выбросов.

Итеративная стохастическая процедура при достаточно большом размере кластера основное вычислительное время тратит на верификацию полученного преобразования на исходных данных. С помощью описанных выше предварительных проверок на согласованность количество итераций можно существенно сократить. Аналитически было установлено, что вероятность ошибочной пары сопоставлений пройти эти проверки не превышает $1/27$. Эта оценка была подтверждена экспериментами.

Таким образом, использование только указанных проверок перед выбором пары сопоставлений аналогично использованию 27 итераций алгоритма RANSAC. В этом случае вероятность P_{cov} получения правильного решения после подбора параметров, удовлетворяющих проверкам, при 50% ошибок:

$$P_{cov}(f=0,5) \geq 1 - [1 - (0,5)^2]^{27} = 0,9996.$$

В результате можно вообще отказаться от проверки гипотезы на всей совокупности сопоставлений из кластера при сохранении очень высокой вероятности получения правильного решения. Таким образом, по сравнению с методом, основанном на МНК, удастся получить выигрыш не только в качестве, но и в скорости работы программы.

Практические результаты

Для того чтобы провести тестирование разработанных алгоритмов, была составлена база данных из 350 изображений (размером 640×480), снятых внутри помещений. 352 пары изображений из этой базы данных были промаркированы вручную. На каждом из двух изображе-

ний были отмечены около 10–15 соответствующих друг другу реперных точек. Поскольку такие сопоставления не содержат ошибок, то по этим точкам можно достаточно точно восстановить фундаментальную матрицу, используя которую, можно проверять, лежат ли сопоставления, найденные автоматически, на соответствующих эпиполярных линиях.

Для сравнения работы разных методов были выбраны следующие критерии:

1. Среднее количество сопоставленных точек.
2. Среднее отклонение сопоставленных точек от эталонных эпиполярных линий.
3. Среднее количество правильно найденных сопоставлений (находящихся от соответствующей эпиполярной линии не дальше $1/20$ от длины диагонали изображения).
4. Количество правильно сопоставленных пар изображений (для них количество правильных сопоставлений должно быть больше 50%).

Результаты для нескольких алгоритмов обнаружения и описания ключевых точек приведены в табл. 2.

По данным таблицы видно, что разработанный метод значительно превосходит метод, основанный на восстановлении эпиполярной геометрии. Также приведены результаты, полученные с помощью платного программного обеспечения – библиотеки ERSP [10], которая до сих пор превосходит любую из программ, находящихся в открытом доступе. Можно заметить, что сочетания разных методов обнаружения и описания ключевых точек дают приблизительно одинаковый результат, что подтверждает то, что ключевую роль в окончательном результате играют алгоритмы сопоставления. Разработанная процедура удаления ошибок (кластеризация и верификация)

Таблица 2. Результаты сравнения разных алгоритмов сопоставления ключевых точек.

Метод детекции	Метод описания	Метод сопоставления	Изобр. сопоставлено	Правильных сопост., %	Среднее отклонение, пикс	Среднее кол-во сопоставл.
SURF	SURF	Эпиполярная геометрия	205	53,76	106,92	285
ERSP Library			286	77	43,77	55
SIFT	SIFT	Кластерный анализ	300	82,85	32,73	56
SURF	SURF	Кластерный анализ	309	83,09	33,1	92
SURF	BRISK	Кластерный анализ	307	82,82	36,4	93
SURF	HOG	Кластерный анализ	309	82,99	34,7	52

занимает в среднем 20 мс на стационарном компьютере при обработке порядка 500 сопоставлений и имеет большой потенциал для оптимизации. Использование разработанных алгоритмов в сочетании с детектором SURF и дескриптором Binary Robust Invariant Scalable Keypoints (BRISK) [11] (реализация из библиотеки OpenCV) обеспечивает возможность обработки 10 кадров в секунду (размером 640×480).

Заключение

Для сопоставления изображений трехмерных сцен предлагается использовать метод, основанный на кластеризации локальных признаков посредством преобразования Хафа. Был разработан способ верификации кластеров сопоставленных ключевых точек, обеспечивающий значительно лучшие результаты, чем методы, использующие метод наименьших квадратов.

Разработанные алгоритмы являются альтернативой использованию фундаментальной матрицы при сопоставлении изображений трехмерных сцен и значительно превосходят этот метод по качеству работы. Предложенный метод характеризуется высокой вычислительной эффективностью и большей функциональностью – по сравнению с эпиполярной геометрией разработанные алгоритмы не накладывают ограничения на жесткую связность расположенных на сцене объектов, и поэтому возможно распознавание динамических сцен и осуществление слежения сразу за несколькими объектами одновременно.

Благодарности

Работа выполнена при поддержке Министерства Образования и Науки Российской Федерации и при государственной финансовой поддержке ведущих университетов Российской Федерации (субсидия 074-U01).

* * * * *

ЛИТЕРАТУРА

1. *Loui A., Das M.* Matching of complex scenes based on constrained clustering // AAAI Fall Symp.: Multimedia Information Extraction. 2008. V. FS-08-05. P. 28–30.
2. *Lutsiv V., Potapov A., Novikova T., Lapina N.* Hierarchical 3D structural matching in the aerospace photographs and indoor scenes // Proc. SPIE. 2005. V. 5807. P. 455–466.
3. *Peterson M.V.* Clustering of a set of identified points on images of dynamic scenes, based on the principle of minimum description length // J. Opt. Techn. 2010. V. 77. P. 701–706.
4. *Potapov A.S., Malyshev I.A., Puysha A.E., Averkin A.N.* New paradigm of learnable computer vision algorithms based on the representational MDL principle // Proc. SPIE. 2010. V. 7696. P. 769606.
5. *Ballard D.H.* Generalizing the Hough transform to detect arbitrary shapes // Pattern Recognition. 1981. V. 13. № 2. P. 111–122.
6. *Lowe D.G.* Object recognition from local scale-invariant features // The Proc. 7th IEEE Intern. Conf. Computer Vision. V. 2. September 20–27 1999. Kerkyra, Greece. P. 1150–1157.
7. *Bay H., Tuytelaars T., Van Gool L.* SURF: Speeded Up Robust Features // Proc. 9th European Conf. on Computer Vision. May 7–13 2006. Graz, Austria. P. 404–417.
8. *Lowe D.* Local feature view clustering for 3D object recognition // IEEE Conf. Computer Vision and Pattern Recognition. December 2001. Kauai, Hawaii, USA. P. 682–688.
9. *Raguram R., Frahm J.M., Pollefeys M.* A comparative analysis of RANSAC techniques leading to adaptive real-time random sample consensus // Proc. European Conf. Computer Vision. October 12–18 2008. Marseille, France. P. 500–513.
10. ERSP 3.1. Robotic Development Platform [Электронный ресурс]. Режим доступа: http://www.mobile-vision-technologies.eu/archiv/download/MVT_ersp.pdf (дата обращения 10.03.2014).
11. *Leutenegger S., Chli M., Siegwart R.* BRISK: Binary Robust Invariant Scalable Keypoints // Proceedings of International Conference on Computer Vision. November 8–11, 2011. Barcelona, Spain. P. 2548–2555.