

Алгоритм сопоставления изображений на основе скользящих гистограмм направленных градиентов

Д. Мирамонте-Харамилло*, В.Х. Диас-Рамирес**, В.И. Кобер***, В.Н. Карнаухов***

* Факультет вычислительной техники, Центр научных исследований и высшего образования, Энсенада, 22860, Мексика

** Национальный политехнический институт - Центр исследований и развития цифровых технологий, Тихуана, 22510, Мексика

*** Институт проблем передачи информации, Российская академия наук, Москва, 127994, Россия

Поступила в редколлегию 17.2.2014

Аннотация—Предложен новый алгоритм сопоставления изображений, основанный на рекурсивном вычислении гистограмм направленных градиентов по нескольким круглым скользящим окнам и пирамидальному разложению изображения. Алгоритм дает хорошие результаты для геометрически искаженных и масштабированных изображений. Результаты компьютерного моделирования, полученные с помощью предложенного алгоритма, сравниваются с точки зрения точности сопоставления и времени обработки с результатами известных алгоритмов.

КЛЮЧЕВЫЕ СЛОВА: сопоставление изображений, быстрый алгоритм, гистограмма направленных градиентов, круглое окно.

ВВЕДЕНИЕ

В последнее время были предложены многочисленные алгоритмы сопоставления изображений, использующие различные признаки и особые точки. Среди них большое внимание привлекают два следующих алгоритма: преобразование признаков, инвариантное к масштабу (SIFT — Scale Invariant Feature Transform) [1] и робастные ускоренные признаки (SURF — Speeded-Up Robust Features) [2]. Эти алгоритмы и их варианты [3]–[5] можно рассматривать в качестве ориентиров для сравнения с предлагаемыми новыми методами сопоставления изображений. Несмотря на то, что базирующиеся на признаках методы очень популярны, тем не менее, алгоритмы сопоставления, базирующиеся на шаблонах, являются притягательной альтернативой для приложений, работающих в реальном времени. Методы, использующие шаблоны, имеют хорошее формальное обоснование. Кроме того, такие методы могут быть реализованы с высокой скоростью обработки в гибридных оптико-цифровых системах [6], в высокопроизводительных аппаратных средствах таких как графические процессоры (GPU) [7] или в программируемых матричных устройствах (FPGA) [8]. Другим подходом решения данной проблемы является сочетание алгоритмов сопоставления, базирующихся как на признаках, так и на шаблонах. Примером может служить инвариантное к масштабу преобразование сжатой гистограммы (SICHT — Scale Invariant Compressed Histogram Transform) [9], в котором в качестве признаков используются гистограммы направленных градиентов [10], вычисляемые в скользящем окне.

Основные достоинства предлагаемого алгоритма:

- Алгоритм выполняет сопоставление гистограмм направленных градиентов в нескольких круглых окнах, скользящих по входному изображению. Окончательное решение принимается на основе результатов совместных сопоставлений для всех окон.

- Для ускорения процесса обработки алгоритм использует прореживание. Локальная пороговая фильтрация гистограмм с последующим прореживанием результирующего изображения используется вместо классического пирамидального подхода с фильтрацией низких частот.
- Возможен компромисс между сложностью алгоритма и точностью сопоставления, контролируемый заданными значениями вероятностей ложных тревог и ошибок пропуска, с одной стороны, и временем обработки, с другой стороны.

Качество работы предлагаемого алгоритма, оцениваемое с использованием тестовой базы данных изображений, сравнивалось с качеством работы SIFT, SURF и более нового ORB (Oriented FAST и Rotated BRIEF) алгоритмов [3] с точки зрения точности сопоставления и времени обработки.

1. АЛГОРИТМ БЫСТРОГО СОПОСТАВЛЕНИЯ ИЗОБРАЖЕНИЙ

Определим некоторый набор круглых окон $\{W_i, i = 1, \dots, M\}$ на эталонном изображении как набор следующих закрытых дисков:

$$W_i = \left\{ (x, y) \in \mathbb{R}^2 : (x - x_i)^2 + (y - y_i)^2 \leq r_i \right\}, \quad (1)$$

где (x_i, y_i) — координаты центра окна и r_i — радиус i -го диска. Диски формируют некоторую геометрическую структуру с относительными расстояниями и углами между центрами окон. Гистограммы направленных градиентов вычисляются в круглых областях и используются для сопоставления. Следует отметить, что в любой позиции структуры каждый диск содержит область изображения, которая не изменяется при повороте. Следовательно, гистограммы ориентированных градиентов, вычисляемые в круглых областях также инвариантны к повороту. Использование только одного окна для сопоставления дает низкие результаты сопоставления. Поэтому рекомендуется выбирать такое минимальное количество дисков, чтобы заполнить как можно большую внутреннюю область эталонного объекта. Гистограммы ориентированных градиентов являются хорошими признаками для сопоставления [10] потому, что они обладают хорошей дискриминантной способностью и являются устойчивыми к небольшим деформациям, таким как поворот и масштабирование.

В каждом положении i -го круглого окна на изображении сцены мы вычисляем градиенты внутри окна с помощью оператора Собеля [11]. Затем, используя модули градиентов $\{Mag_i(x, y) : (x, y) \in W_i\}$ и величины углов направлений, квантованные на Q уровней $\{\varphi_i(x, y) : (x, y) \in W_i\}$, вычисляем гистограммы направленных градиентов:

$$HoG_i(\alpha) = \begin{cases} \sum_{(x,y) \in W_i} \delta(\alpha - \varphi_i(x, y)), & \text{если } Mag_i(x, y) \geq Med, \\ 0, & \text{в противном случае,} \end{cases} \quad (2)$$

где $\alpha = \{0, \dots, Q - 1\}$ — номер дискретного уровня гистограммы, Med — медиана модулей градиентов круглых окон и $\delta(z) = \begin{cases} 1, & \text{если } z = 0, \\ 0, & \text{в противном случае} \end{cases}$ — дельта-функция Кронекера.

Отметим, что вычисления в уравнении (2) требуют приблизительно $\lceil \pi r_i^2 \rceil$ операций сложения, где $\lceil \cdot \rceil$ — оператор целой части вещественного числа. Для снижения вычислительной сложности, вычисление гистограммы в любой позиции скользящего окна может быть выполнено рекурсивным способом. Такой подход приводит к вычислительной сложности порядка $\lceil 2\pi r_i \rceil$ операций сложения. Рис. 1 показывает рекурсивное вычисление гистограммы вдоль столбцов

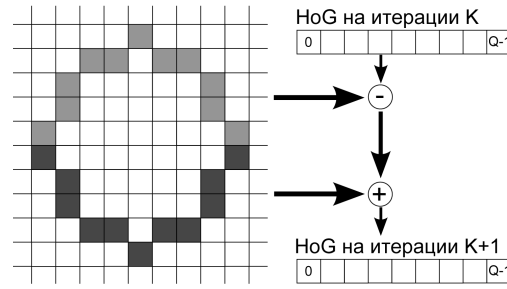


Рис. 1. Рекуррентное изменение гистограммы вдоль столбцов при переходе от итерации k к итерации $k + 1$.

при переходе от итерации k к итерации $k + 1$. Для обеспечения инвариантности к повороту мы используем нормализованный корреляционный оператор для сравнения гистограмм эталонного изображения и сцены. Вычислим центрированную и нормализованную гистограмму направленных градиентов эталонного изображения в следующем виде:

$$\overline{HoG_i^R}(\alpha) = \frac{HoG_i^R(\alpha) - Mean^R}{\sqrt{Var^R}}, \quad (3)$$

где $Mean^R$ и Var^R — выборочное среднее значение и дисперсия гистограммы соответственно.

Корреляционная функция для каждого i -го окна в позиции k может быть вычислена с помощью обратного преобразования Фурье [11] как:

$$C_i^k(\alpha) = IFT \left[\frac{HS_i^k(\omega) HR_i^*(\omega)}{\sqrt{Q \sum_{q=0}^{Q-1} (HoG_i^k(q))^2 - (HS_i^k(0))^2}} \right], \quad (4)$$

где $HS_i^k(\omega)$ — преобразование Фурье гистограммы направленных градиентов внутри i -го окна входной сцены, $HR_i(\omega)$ — преобразование Фурье $\overline{HoG_i^R}(\alpha)$, а звездочка $*$ обозначает комплексное сопряжение. Корреляционный пик является мерой схожести двух гистограмм и может быть получен как:

$$P_i^k = \max_{\alpha} \{C_i^k(\alpha)\}. \quad (5)$$

Значения этих корреляционных пиков находятся в диапазоне $[-1, 1]$. Следует подчеркнуть, что нормализованные корреляционные пики (5) обладают двумя важными свойствами. Первый из них — инвариантность к повороту поскольку циклический сдвиг значений гистограммы соответствует циклическому сдвигу значений пиков. Вторым — нормализация в уравнениях (3) и (4) позволяет учесть небольшую разницу между масштабами эталонного изображения и сцены. Вычисление центрированных и нормализованных гистограмм для всех круглых окон по эталонному изображению и вычисление преобразований Фурье могут быть выполнены как предварительная обработка эталонных данных.

Для ускорения процесса сопоставления мы не используем классический пирамидальный подход, который основан на низкочастотной фильтрации и последующей дискретизации результирующего изображения для получения изображений низкого разрешения меньшего размера [12]. Простое прореживание входного изображения вдоль строк и столбцов [13] представляется более подходящим для ускорения процесса сопоставления гистограмм направленных градиентов.

Сопоставление выполняется между первым уменьшенным изображением сцены и уменьшенным эталонным изображением. Заметим, что разложение эталонного изображения на множество уменьшенных изображений может быть выполнено как предварительная обработка эталонных данных. Другим важным параметром, который может ускорить процесс сопоставления, является количество дискретных уровней гистограмм Q .

Предлагаемая процедура сопоставления кратко может быть описана в следующем виде:

- Выполняется независимое быстрое сопоставление для круглых окон с уменьшенным числом дискретных уровней гистограмм, скажем Q_R . Осуществляется уменьшение числа точек, которые будут рассматриваться на следующем проходе сопоставления при сохранении низкого значения вероятности ошибок пропуска.
- Только оставшиеся выделенные точки рассматриваются для выполнения второго прохода сопоставления при стандартном количестве дискретных уровней гистограмм Q для каждого окна.
- Дальнейшее снижение как вероятности ошибок пропуска, так и вероятности ошибок ложной тревоги достигается на основе совместного анализа признаков во всех рассматриваемых круглых областях эталона.

Детальное описание основных блоков алгоритма и пояснения к ним могут быть обобщены следующим образом:

1. Определить на эталонном изображении набор круглых окон $\{W_i, i = 1, \dots, M\}$ постоянного радиуса r , используя уравнение (1). Вычислить расстояние между центрами окон

$$D_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}, \quad i = 1, \dots, M; \quad j = i + 1, \dots, M$$

и углы между каждыми тремя смежными центрами всей совокупности окон, т.е.

$$\gamma_i = \cos^{-1} \left[\frac{D_{i,j+1}^2 + D_{i,j+2}^2 - D_{i,j+3}^2}{2D_{i,j+1}D_{i,j+2}} \right], \quad i = 1, \dots, M - 2.$$

Установить такое пороговое значение Th_{Q_R} , которое дает заданное значение вероятности ошибок пропуска при уменьшенном числе дискретных уровней гистограмм Q_R . Подобным способом установить пороговое значение Th_Q , которое обеспечит компромисс между вероятностями пропуска и вероятностями ложной тревоги для Q дискретных уровней гистограмм.

2. Проредить эталонное изображение с шагом, равным L , вдоль столбцов и строк и получить L^2 уменьшенных изображений. Для каждого круглого окна эталонного изображения вычислить, используя уравнения (2) и (3), центрированные и нормализованные гистограммы направленных градиентов $\{HoG_i^R(\alpha), i = 1, \dots, M\}$ и вычислить преобразование Фурье $\{HR_i(\omega), i = 1, \dots, M\}$.
3. Выполнить сжатие $\{HoG_i^R(\alpha), i = 1, \dots, M\}$ при количестве дискретных уровней гистограмм, равном Q , и вычислить преобразование Фурье $\{CHR_i(\omega), i = 1, \dots, M\}$.
4. Проредить изображение сцены с шагом, равным L , вдоль столбцов и строк и получить L^2 уменьшенных изображений. Начать процесс сопоставления при использовании всех круглых окон изображения сцены для $k = 1$. Поскольку все окна имеют один размер, то используется только одно скользящее круглое окно. Индекс i гистограммы изображения сцены может быть опущен. Вычислить гистограмму направленных градиентов первого уменьшенного изображения сцены $HoG^1(\alpha)$ и его преобразование Фурье $HS^1(\omega)$. Выполнить сжатие

- $Hog^1(\alpha)$ при количестве дискретных уровней гистограмм, равном Q , для того чтобы получить $CHog^1(\alpha)$ при количестве дискретных уровней гистограмм, равном Q_R , и вычислить преобразование Фурье $CHS^1(\omega)$.
5. Вычислить корреляционные пики между $CHog^1(\alpha)$ и $\{\overline{CHog_i^R}(\alpha), i = 1, \dots, M\}$, используя уравнения (4) и (5).
 6. Если любой из корреляционных пиков, вычисленных на шаге 5, больше, чем Th_{QR} , то вычислить корреляционные пики $\{P_i^1, i = 1, \dots, M\}$ между $Hog^1(\alpha)$ и $\{\overline{Hog_i^R}(\alpha), i = 1, \dots, M\}$, используя уравнения (4) и (5). Если любой из корреляционных пиков $\{P_i^1, i = 1, \dots, M\}$ больше чем Th_Q , то сохранить эту точку для дальнейшего анализа.
 7. Увеличить k и обновить гистограммы направленных градиентов изображения сцены вдоль столбцов и строк и повторить шаги 5 и 6.
 8. Принять окончательное решение о наличии эталонного объекта в позиции k . Во время анализа мы рассматриваем возможную деформацию исходной структуры, обусловленную масштабированием и поворотом сцены вне плоскости изображения. Начало анализа начинается в позиции самого высокого пика, вычисленного для первого окна. Затем, определяется кольцевая область с внешним и внутренним радиусами, равными $D_{1,2} \pm \epsilon_D$ соответственно, для поиска пика внутри второго окна. Поиск последовательных корреляционных пиков может быть выполнен быстро при учете углов между центрами окон. Например, для третьего окна мы вычисляем пересечение между кольцевой областью с внешним и внутренним радиусами $D_{2,3} \pm \epsilon_D$, и открытым круговым сектором с началом в позиции пика второго окна, и ограниченными внутренним и внешним радиусами с наклонами $\gamma_1 \pm \epsilon_\gamma$.

2. РЕЗУЛЬТАТЫ ЭКСПЕРИМЕНТОВ

В этом разделе мы представляем результаты экспериментов, выполненных с использованием базы данных изображений ALOI [14]. Были использованы десять изображений сцен размером 1280×1024 пикселей и десять эталонных изображений размером 144×144 пикселей с различными эталонными объектами. Каждый эталонный объект располагался случайным образом на сценах в 100 различных позициях. Качество предложенного алгоритма сравнивалось с качеством алгоритмов SIFT, SURF и более нового — ORB [3]. Параметры первых двух алгоритмов были взяты из соответствующих статей [1] и [2]. Параметры для ORB алгоритма были эмпирически выбраны таким образом, чтобы получить сопоставимые по качеству и времени обработки результаты. Алгоритмы были протестированы в различных условиях, таких как повороты изображений в плоскости/вне плоскости сцены и небольшие изменения масштаба. Качество оценивалось с точки зрения количества правильных сопоставлений и времени обработки. Предложенный алгоритм, который будем называть как Circular Window – Histograms of Oriented Gradients (CW-HOG), использует два круглых окна с радиусом r (зависит от размера объекта). Для получения лучшего качества сопоставления мы используем $Q_R = 16$ and $Q = 64$ дискретных уровней гистограмм вместо 9 уровней, предложенных в [10]. Параметры алгоритма в экспериментах имеют следующие значения:

- $M = 2, Q = 64, Q_R = 16, \epsilon_D = 0.2r, L = 4$;
- Пороги для принятия решений заданы следующими: при $Q_R = 16$ (первый проход) порог $Th_{QR} = 0.8$ дает вероятность ошибки пропуска на тестовой базе данных $P_{ME} = 0.05$ и выделяет приблизительно 46 точек для их дальнейшего анализа на втором проходе при $Q = 64$;
- При $Q = 64$ (второй проход) порог $Th_Q = 0.7$ обеспечивает вероятность ошибки пропуска $P_{ME} = 0.05$ и вероятности ложного обнаружения $P_{FA} = 0.21$;

– Окончательное решение снижает вероятность ложной тревоги до $P_{FA} = 0.05$.

Точность сопоставления изображений при использовании выбранных алгоритмов в зависимости от угла поворота изображения в плоскости изображения сцены показана на Рис. 2. Нетрудно видеть, что CW-HOG алгоритм дает сопоставимое качество, с точки зрения инва-

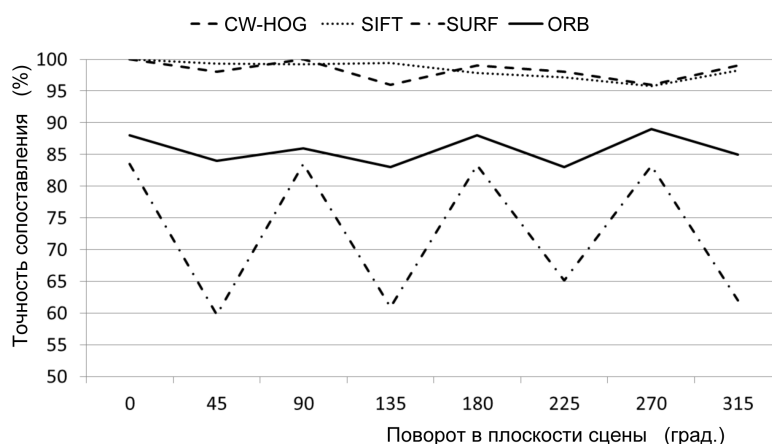


Рис. 2. Точность сопоставления изображений при использовании выбранных алгоритмов в зависимости от угла поворота изображения вне плоскости изображения сцены.

риантности в повороту, в сравнении с SIFT. Точность сопоставления изображений при использовании выбранных алгоритмов в зависимости от угла поворота изображения вне плоскости изображения сцены показана на Рис. 3. Отметим, что CW-HOG алгоритм дает более

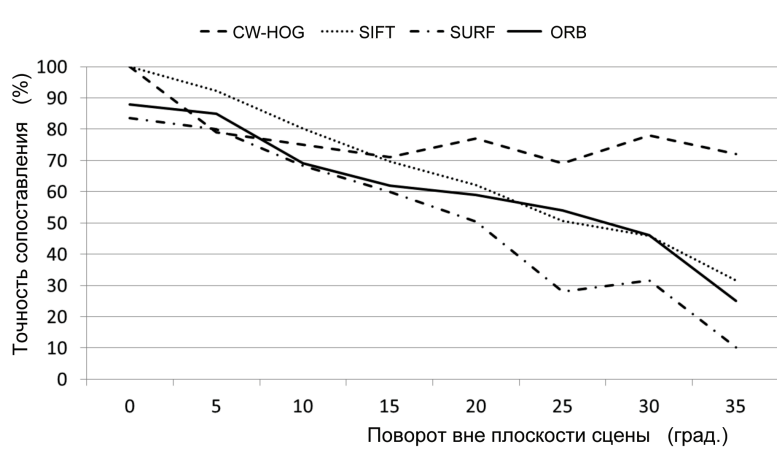


Рис. 3. Точность сопоставления при использовании выбранных алгоритмов в зависимости от угла поворота изображения вне плоскости изображения сцены.

стабильное качество сопоставления изображений при поворотах изображения вне плоскости изображения сцены в сравнении с другими алгоритмами. Рис. 4 иллюстрирует устойчивость предложенного алгоритма к изменению масштаба сцены в диапазоне $[0.8, 1.2]$. Заметим, что точность сопоставления изображений алгоритмом CW-HOG является довольно хорошей и в этом случае.

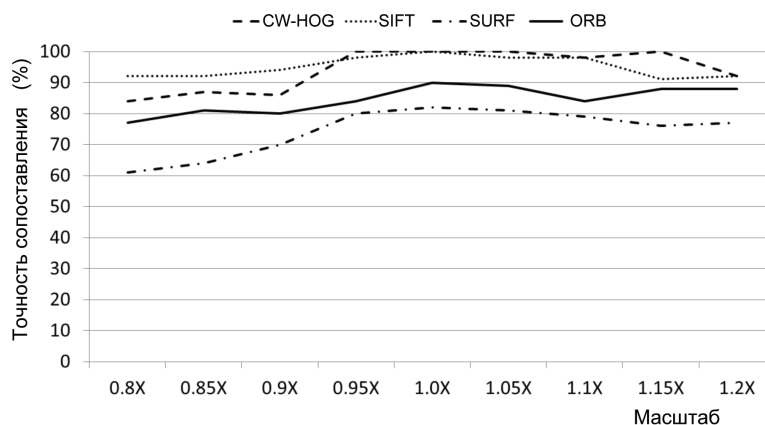


Рис. 4. Точность сопоставления при использовании выбранных алгоритмов в зависимости от небольших изменений масштаба изображения сцены.

Сравнение выбранных алгоритмов, с точки зрения времени обработки, показало, что время работы алгоритмов CW-HOG и ORB составляет порядка 1.7 с, а время работы алгоритмов SIFT и SURF составляет 9.82 с и 0.95 с соответственно. Алгоритмы были реализованы на стандартном персональном компьютере с процессором Intel Core i7 с тактовой частотой 3,2 ГГц и оперативной памятью 8 Гб. Реализация алгоритмов SIFT, SURF и ORB была взята из открытой библиотеки OpenCV с библиотекой Intel multithreading library TBB. Предложенный алгоритм CW-HOG был также реализован на основе открытой библиотеки OpenCV с библиотекой multithreading library OpenMP.

Алгоритмы слежения за объектами обычно используют методы предсказания [15] для оценки положения и траектории объекта в последовательности кадров и для уменьшения области поиска. Характеристики предложенного алгоритма, с точки зрения времени обработки от размера фрагментов сцены, следующие:

- фрагменты 640×512 пикселей — 0.47 с;
- фрагменты 426×341 пикселей — 0.22 с;
- фрагменты 320×256 пикселей — 0.14 с.

Очевидно, что время обработки резко падает при уменьшении размера фрагмента. Следовательно, становится возможным реализация слежения за объектом на основе предложенного алгоритма.

ВЫВОДЫ

В работе представлен быстрый алгоритм сопоставления изображений, основанный на рекурсивном вычислении гистограмм направленных градиентов по нескольким круглым скользящим окнам. Мультиядерные процессоры с интегрированной параллельной архитектурой могут быть использованы для реализации предложенного алгоритма сопоставления изображений большого размера с высокой скоростью. Представленные результаты экспериментов показали, что предложенный алгоритм превосходит известные алгоритмы при повороте изображения в плоскости изображения сцены, дает подобные с алгоритмом SIFT результаты при повороте изображения вне плоскости изображения сцены и обладает скоростью обработки близкой к алгоритму SURF. Предложенный алгоритм является перспективным для таких приложений реального времени, как слежение за объектами, когда требуется выполнять быстрое сопоставление изображений, инвариантное к повороту и небольшому изменению масштаба.

СПИСОК ЛИТЕРАТУРЫ

1. Lowe D.G. Object recognition from local scale-invariant features. *Proc. International Conference on Computer Vision*, 1999, vol. 2, pp. 1150–1157.
2. Bay H., Ess A., Tuytelaars T., Van Gool L. SURF: Speeded Up Robust Features. *Computer Vision Image Understanding*, 2008, vol. 110, no. 3, pp. 346–359.
3. Rublee E., Rabaud V., Konolige K., Bradski G. ORB: an efficient alternative to SIFT or SURF. *IEEE International Conference on Computer Vision*, 2011, pp. 2564–2571.
4. Calonder M., Lepetit V., Strecha C., Fua P., BRIEF: binary robust independent elementary features. *Proc. 11th European Conference on Computer Vision*, 2010, ECCV'10, pp. 778–792.
5. Ortiz R. FREAK: Fast Retina Keypoint. *Proc. of the 2012 IEEE Conference on Computer Vision and Pattern Recognition*, CVPR'12, 2012, pp. 510–517.
6. Díaz-Ramírez V.H., Kober V., Adaptive phase-input joint transform correlator. *Applied Optics*, 2007, vol. 46, no. 26, pp. 6543–6551.
7. Ouerhani Y., Jridi M., Alfalou A., Brosseau C. Optimized preprocessing input plane GPU implementation of an optical face recognition technique using a segmented phase only composite filter. *Optical Communications*, 2013, vol. 289, pp. 33–44.
8. Rice K.L., Taha T.M., Chowdhury A.M., Awwal A.A.S., Woodard D.L. Design and acceleration of phase-only filter-based optical pattern recognition for fingerprint identification. *Optical Engineering*, 2009, vol. 48, no. 11, pp. 117–206.
9. Zalesky B.A., Lukashovich P.V. Scale Invariant Algorithm to Match Regions on Aero or Satellite Images. *Proc. Pattern Recognition and Information Processing*, 2011, vol. 11, pp. 25–30.
10. Dalal N., Triggs B. Histograms of Oriented Gradients for Human Detection. *Computer Vision and Pattern Recognition*, 2005, vol. 1, pp. 886–893.
11. Pratt W.K. *Digital Image Processing*, John Wiley & Sons, 2007.
12. Lindeberg T. Scale-space theory: A basic tool for analysis structures at different scales. *Journal of Applied Statistics*, 1994, vol. 21, no. 2, pp. 225–270.
13. Liu B., Zaccarin A., New fast algorithms for the estimation of block motion vectors. *IEEE Transactions on Circuits and Systems for Video Technology*, 1993, vol. 3, no. 2, pp. 148–157.
14. Geusebroek J.M., Burghouts G.J., Smeulders A.W.M. The Amsterdam library of object images. *International Journal of Computer Vision*, 2005, vol. 61, no. 1, pp. 103–112, <http://staff.science.uva.nl/~aloi/>
15. Rong Li X., Jilkov V.P. Survey of maneuvering target tracking. Part I: Dynamic models. *IEEE Transactions on Aerospace and Electronic Systems*, 2003, vol. 39, no. 4, pp. 1333–1364.

A novel image matching algorithm based on sliding histograms of oriented gradients

Miramontes-Jaramillo D., Díaz-Ramírez V.H., Kober V., Karnaukhov V.

A new image matching algorithm based on recursive calculation of histograms of oriented gradients over several circular sliding windows and pyramidal image decomposition is presented. The algorithm yields a good performance for geometrically distorted and scaled scene images. Computer results obtained with the proposed algorithm are compared with those of common algorithms in terms of matching accuracy and processing time.

KEYWORDS: image matching, fast algorithm, histogram of oriented gradients, circular window.