

Методы оценки и подготовки изображений в видеопотоке к распознаванию объектов

С.Г. Небаба¹
stepanlfx@tpu.ru

¹Томский политехнический университет, Томск, Россия

В работе рассматривается проблема выбора методов оценки и подготовки изображений в видеопотоке для систем распознавания объектов реального времени. Рассмотрены основные способы компенсации условий съемки и оптимизации подготовки изображений к распознаванию объектов, такие как оценка контрастности изображений, резкости, нормализация освещения, поиск и отбраковка похожих изображений. Проведен анализ существующих алгоритмов сравнения изображений, оценки контрастности и резкости изображения.

Ключевые слова: компьютерное зрение, распознавание объектов, яркость, контрастность, резкость, видеопоток, режим реального времени.

Methods of evaluation and preparation of images in the video stream for object recognition

S.G. Nebaba¹
stepanlfx@tpu.ru

¹Tomsk Polytechnic University, Tomsk, Russia

The problem of choosing methods for estimating and preparing images in the video stream for real-time object recognition systems is considered in this paper. The main methods for compensating shooting conditions and optimizing the preparation of images for object recognition, such as estimating the contrast of images, sharpening, lighting normalization, searching and rejecting similar images, are considered. The current estimators of similar images search, the contrasting effect and image sharpness analyzed.

Keywords: computer vision, object recognition, brightness, contrast, sharpness, video stream, real-time mode.

1. Введение

Дисциплины компьютерного зрения и робототехники являются одними из наиболее активно развивающихся областей науки. Одной из фундаментальных проблем в этих сферах является оценка и нормирования уровней яркости, контрастности и цветности изображений. Эффективное решение этих проблем упрощает задачи слежения за объектами, их идентификации, определения и моделирования их свойств и внешнего вида [1].

С развитием информационных технологий, средств коммуникации, мобильных устройств и робототехнических комплексов наметился бурный прогресс в решении такого рода задач. Однако на сегодняшний день все еще не существует универсального способа быстрой и надежной идентификации объектов, даже с привлечением 3D моделей высокого разрешения на основе систем, требующих обработки данных в режиме реального времени.

Системы технического зрения (СТЗ) являются основной частью многих интеллектуальных систем, а также управляемых и автономных роботов. Сложность задач, решаемых такими системами, постоянно увеличивается, а требования к их техническим характеристикам повышаются. При этом все острее встают вопросы надежности и точности подобных автоматизированных систем, особенно для таких критичных направлений как безопасность предприятий, автономные транспортные средства, мобильные роботы и медицинские технологии. Продолжает расти потребность в быстрой некооперативной идентификации личности по изображению лица, в особенности в сфере безопасности, где оперативность реакции на событие имеет решающее значение [2].

В связи с этим возникает необходимость систематизации существующих и разработке новых методов и алгоритмов высокоскоростной оценки и обработки изображений. В перспективе это способствует

повышению точности распознавания объектов в видеопотоке в режиме реального времени по его изображению, в том числе для автономных мобильных систем с возможностью массового применения и реализации.

В работах [3,4] была предложена комплексная технология обработки и подготовки изображений лиц в видеопотоке к распознаванию личности, рассчитанная на работу в режиме реального времени (25 кадров в секунду). Предложенная технология позволила повысить точность распознавания личности по изображению лица по сравнению с математическими классификаторами и добиться высокой скорости обработки видеопотока, превосходя по производительности нейросетевые методы распознавания. Однако, несмотря на эффективность предложенной технологии и повышение точности распознавания личности в режиме реального времени, проблема выбора методов оценки и нормализации освещения сцены, яркости, контрастности и цветности изображений в видеопотоке была рассмотрена недостаточно подробно и существует потенциал для улучшения системы в этом направлении.

Таким образом, возникает задача исследования методов, а также создания системы, которая будет способна осуществлять подбор наиболее эффективных методов оценки и нормирования изображений для конкретных задач распознавания объектов с сохранением высокой скорости обработки видеопотока.

2. Информативность изображения в видеопотоке

Для систем распознавания объектов чаще всего существуют неформализованные требования, что условия освещения должны позволять регистрировать изображения объектов, на которых могут быть различим основной рельеф (достаточный контраст). Важным условием

является наличие прямого или близкого к нему угла падения лучей света на фронтальную плоскость объектов, а также наличие рассеянного источника света. Заметное отрицательное влияние на качество распознавания оказывает боковая засветка объектов.

В большей части систем видеонаблюдения используется частота поступления кадров не менее 25 кадров в секунду. Очевидно, что изображения анализируемого объекта на двух соседних кадрах будут почти неотличимы друг от друга. Анализ таких практически идентичных кадров не оправдан с точки зрения использования вычислительных ресурсов. В связи с этим имеет смысл проводить отбраковку мало отличимых кадров.

Предварительный обзор позволил выделить 3 основных подхода сравнения изображений, получивших широкое распространение.

1. Применение хэш-функций к уменьшенной копии изображения и их сравнение. Данный метод обладает хорошим быстродействием, но дает очень грубую оценку, которая напрямую зависит от уменьшения изображения, поэтому может применяться только для поиска отдаленного визуального сходства объектов, например, поиску похожих картинок в коллекции. Метод относится к поиску изображений по содержанию (англ. Content-based image retrieval (CBIR)) [5] – разделу компьютерного зрения, решающему задачу поиска изображений, которые имеют требуемое содержание, в большом наборе цифровых изображений.

2. Вычисление коэффициента корреляции по одной из формул. Этот метод достаточно чувствителен к небольшим изменениям изображения, однако требователен к вычислительным ресурсам. Его применение ограничено специфическими областями работы с изображениями [6].

3. Сравнение изображений через построение и анализ SURF-дескрипторов, которые описывают некоторые небольшие области изображения. Исследования показывают, что данный метод показывает наилучшие результаты при сравнении изображений объектов, снятых под разными углами и в разных условиях освещения, однако он требует в несколько раз больше вычислительных ресурсов, чем сравнительно трудоемкий метод вычисления коэффициента корреляции [7]. Существует модификация данного метода (BRISK), которая обладает высоким быстродействием [8], однако данный метод сложно применить к поиску малозаметных отличий в кадрах.

Таким образом, задача сравнения кадров не имеет простого и эффективного решения, применимого к видеопотоку высокого разрешения, однако остается возможность разработать некий гибридный метод на основе рассмотренных выше.

3. Оценка контрастности

Перед тем, как улучшать изображение, необходимо дать оценку его качеству. Человек, бросив один взгляд на изображение, может сказать яркое оно или темное, четкое или размытое, контрастное или нет и т.д. Алгоритмы же работают детально, анализируя изображение попиксельно или небольшими группами пикселей. Поэтому, на основании работы алгоритма, тяжело дать общую оценку качеству изображения.

Контраст (фр. *contraste*) – в оптике (сенситометрии и фотометрии) разница в характеристиках различных участков изображения, способность фотографического материала или оптической системы воспроизводить эту разницу, а также характеристика чувствительности глаза (зрительной системы) относительно яркости и цвета.

Контрастность (также, в различных контекстах употребляется и само слово контраст и коэффициент контраста) – степень контраста, чаще всего выражается безразмерной величиной, отношением или логарифмом отношений.

Мерой контрастности является величина $k = \frac{B_b - B_s}{B_s}$,

где B_b – яркость фона, B_s – яркость предмета или изображения.

Контрастность изображения является показателем его информативности, так как этот показатель учитывает освещенность и различимость объектов на изображении. Этот параметр недостаточен, чтобы можно было признать изображение качественным, так как боковая засветка объектов дает высокую оценку контраста, но затрудняет выделение деталей изображения.

В случае, если изображение удастся нормализовать, контрастность становится более полезной оценкой.

Различают 2 типа контрастности: яркостная и тоновая.

Яркостная контрастность представляет собой разницу между физической или видимой яркостью отдельных участков изображения. Вообще говоря, вычисление физической или видимой яркости можно рассматривать как конвертацию цветного изображения в ахроматические цвета. Поэтому яркостная контрастность – это сравнение двух участков изображения, приведенных к ахроматическим цветам.

Если проанализировать RGB-гистограммы, то можно сделать вывод, что у контрастного изображения количество темных и светлых пикселей должно быть приблизительно равным, разница в их яркости – значительна, а основное место сосредоточения пикселей – возле границ диапазона.

Тоновая контрастность учитывает не только яркостные переходы, но и переходы цвета. Ее имеет смысл применять при анализе цветных изображений, так как, несмотря на возрастание сложности вычислений, итоговая оценка становится более точной.

Подходящим критерием оценки яркостной контрастности может быть дисперсия яркости пикселей изображения [9]:

$$\sigma^2 = \frac{1}{N} \sum_{p=1}^N (Y_p - Y)^2, \quad (1)$$

где Y – среднее значение яркостной контрастности всего изображения, Y_p – значение яркостной контрастности в точке p , N – общее число точек изображения.

Оценка контрастности (1) может быть нормирована путём вычисления отношения среднеквадратического отклонения к максимально возможному значению яркости:

$$C = \frac{2\sigma}{Y_{\max}}. \quad (2)$$

Диапазон изменения значения C – [0;1]. Значение 0 соответствует однотонному изображению, значение 1 – максимально контрастному. Оптимальное значение контрастности (2) зависит от типа объекта, представленного на изображении.

Контрастность изображения не может являться точным показателем качества изображения: размытые изображения менее информативные, но при этом имеют более высокую оценку контрастности, чем четкие изображения с меньшим числом точек максимальной и минимальной яркости.

4. Оценка резкости

Резкость чаще всего трактуется как качество воспроизведения на фото достаточно крупных фрагментов. В этом, условно, ее отличие от четкости изображения, характеризующей различимость мелких деталей.

Повышение резкости – один из самых распространенных способов улучшения фотографий. Обычно всегда требуется в той или иной степени усиливать резкость первоначального изображения.

Объекты на изображении отличаются друг от друга по яркости или оптической плотности, что является причиной их различимости. Граница же между объектами всегда имеет ту или иную степень размытости.

Ширина размытости теоретически резкой границы между светлыми (яркими) и темными участками (объектами) изображения характеризует резкость изображения. Причем это характерно для любых изображений, как цветных, так и черно-белых.

Четкость изображения – это параметр, который характеризует качество полученного изображения и показывает, насколько отчетливо различимы мелкие объекты на изображении и определяется минимальным расстоянием между ними.

Другими словами, четкость показывает полноту воспроизведения мелких объектов на изображении и насколько у них резкие контуры.

Изображение высокой четкости имеет резкие, четко видимые границы.

Резкость может выступать еще одним критерием информативности изображения, она в общем случае может быть определена наличием контуров контрастных переходов. Однозначного метода для определения резкости не существует, но есть несколько алгоритмов, которые позволяют вычислить оценку резкости из различных характеристик изображения.

Главное различие всех предлагаемых методов состоит в формулах, которые определяют количественную оценку контрастности изображения. В целом все методы можно разделить на 3 группы:

1. методы, вычисляющие среднее значение дифференциальных операторов изображения;
2. методы, вычисляющие эксцесс амплитудного спектра изображения;
3. методы, вычисляющие отношение значимых и незначимых точек амплитудного спектра изображения.

Первая группа методов вычисляет разностную характеристику напрямую из исходного изображения, например с помощью оператора Лапласа или оператора Робертса, вторая и третья группы вычисляют частотные характеристики изображения (преобразование Фурье, либо Wavelet-преобразование), и вычисляет количество контрастных переходов из спектра изображения.

Идея метода, использующего дифференциальные операторы, заключается в вычислении среднего значения яркости от преобразованного оператором изображения. Такой метод чувствителен к качеству изображения, а также к характеру изображенных объектов, то есть он может дать высокую оценку резкости для плохого изображения со сложным объектом, и меньшую – для хорошего изображения с малым числом контрастных переходов.

Оператор Лапласа – дифференциальный оператор, сумма вторых частных производных по координатам изображения, эквивалентен операциям градиента и дивергенции:

$$\text{Laplace}(f) = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$$

Оператор Робертса – свертка изображения с ядрами

$$\begin{bmatrix} +1 & 0 \\ 0 & -1 \end{bmatrix} \text{ и } \begin{bmatrix} 0 & +1 \\ -1 & 0 \end{bmatrix}.$$

Оператор Робертса по сравнению с оператором Лапласа требует меньше вычислений, более точно выделяет тонкие контуры, но очень чувствителен к шуму.

В работе [10] рассмотрен метод оценки резкости на основе преобразования Фурье, суть которого заключается в вычислении эксцесса изображения в частотной области:

$$\gamma_2 = \frac{\mu_4}{\mu_2^2} = \frac{\mu_4}{\sigma^4}, \quad (3)$$

где μ_2 – второй центральный момент изображения, μ_4 – четвертый центральный момент, σ – стандартное отклонение.

В работе [11] также используется преобразование изображения в частотную область, но в качестве критерия оценки резкости предлагается отношение

$$FM = \frac{T_H}{M \cdot N}, \quad (4)$$

где T_H – число значимых точек, $M \cdot N$ – размер изображения амплитудного спектра. Значимые точки определяются как точки, значение в которых превышает некоторую долю максимума амплитуды спектра.

Исследование методов показало, что оценка резкости с использованием оператора Лапласа даёт заметный разброс результатов при схожих условиях освещения и ракурса, что связано отсутствием нормирования результата и способом выделения области объекта. Оценка (3) показывает более стабильные результаты, но практически не реагирует на фактическое снижение резкости, полученное размытием по Гауссу. Оценка (4) наиболее точно соответствует изначальной задаче выделения поиска резких изображений, она корректно реагирует на изменения резкости при движении и отсутствие фокуса, и помимо этого имеет строгую нормировку, позволяющую по заданному уровню отсеивать изображения неудовлетворительного качества.

Несмотря на достоинства оценки (4), она отчасти подвержена тем же проблемам, что и оценка с использованием Лапласиана. Основная причина этого в том, что резкость изображения оценивается количественно, и потому связана с общим числом контрастных переходов, а это число зависит от объекта на изображении. Исходя из этого, можно говорить только об относительной резкости изображения, которая оценивается для одного объекта в треке (серии изображений).

5. Компенсация освещения

Повысить применимость методов (1)-(4) может компенсация изображения объекта по условиям освещения.

Для нормирования исходного изображения по условиям освещения существует несколько методов, например DoG-фильтрация [12] и частотная фильтрация полосовым фильтром [13], убирающим низкие и высокие частоты. Оба метода позволяют эффективно выравнивать изображение при дисбалансе яркости по освещению.

DoG обнаруживает края путем применения размытия изображения по Гауссу (Gaussian blur) с указанным значением стандартного отклонения. Результирующее изображение является размытой версией исходного. Далее применяется другое размытие с меньшим значением стандартного отклонения, что дает более резкое изображение, чем в предыдущем случае. Финальное изображение вычисляется заменой каждого пикселя разностью от пикселей двух размытых изображений.

Для нормализации яркости изображения с этим методом часто используется гамма-коррекция.

Частотный полосовой фильтр при нормализации освещения подавляет высокочастотные шумы и сглаживает неравномерное освещение изображений, кодируемое в изображении низкими частотами.

Оба метода повышают четкость изображения, но не устраняют полностью влияние неравномерного освещения.

Также для повышения контраста изображения применяется адаптивное выравнивание гистограммы с ограничением (CLAHE) [14]. Первоначально изображение разбивается на прямоугольные области (блоки). В каждом блоке вычисляется гистограмма распределения яркости. Чтобы избежать перенасыщения монотонных участков изображения, ограничивается наклон функции распределения яркости путем введения ограничения на максимальное значение частоты яркости. Пусть M – число строк в блоке, N – число столбцов в блоке, α – параметр, задающий ограничение. Предельное значение гистограммы вычисляется в соответствии с формулой:

$$hist_{lim} = hist_{min} + \alpha(MN - hist_{min}),$$

где $hist_{min} = MN/(L+1)$, L – максимальное значение яркости.

После ограничения для каждой области формируется своя функция преобразования яркости на основании выравнивания гистограммы распределения яркости. Для входного изображения вычисляется значение яркостей выходного изображения в соответствии с формулой:

$$LUT(k) = \sum_{j=0}^k p_r(r_j) = \sum_{j=0}^k \frac{n_j}{n},$$

где $LUT(k)$ – значение яркости выходного изображения, соответствующее яркости r_k входного изображения, $k \in [0, L]$; r_j – яркость элемента входного изображения, n_j – число элементов входного изображения со значением r_j , n – общее число элементов изображения; $p_r(r_j)$ – вероятность появления элемента входного изображения с яркостью r_j .

Такое преобразование позволяет равномернее распределить уровни яркости по частоте, в основе преобразования лежит предположение, что в контрастных черно-белых изображениях значения яркости распределены равномерно.

Чтобы разрушить блочную структуру полученного изображения, выходное изображение формируется путем билинейной интерполяции по четырем ближайшим LUT .

6. Заключение

В работе проведен обзор существующих методов подготовки изображений к процессу распознавания объектов и критериев оценки качества изображения.

Проведен анализ существующих алгоритмов сравнения похожих изображений, а также алгоритмов оценки контраста, которые подходят для проверки условий освещения, позволяя предотвратить использование системы распознавания объектов на мало-контрастных изображениях и, как следствие, снизить вероятность ложноположительных срабатываний.

Проведены обзор и сравнение алгоритмов оценки резкости изображения. Оценка резкости позволяет ранжировать изображения в видеопотоке по количеству контрастных переходов, в общем случае давая возможность выбрать из всего количества кадров более качественные.

Рассмотрен метод компенсации условий освещения для изображений с неравномерной яркостью.

Применение всех рассмотренных методов и алгоритмов в подходящих условиях позволит снизить ресурсоемкость систем распознавания объектов по изображению, а также устранить влияние неблагоприятных для процесса распознавания факторов.

7. Благодарности

Работа выполнена в рамках госзадания № 2.1642.2017/4.6 на выполнение проекта по теме «Когнитивные методы визуализации и анализа

многомерных данных при моделировании нелинейных динамических систем».

8. Литература

- [1] Hartley R., Zisserman A. Multiple view geometry in computer vision. – Cambridge university press, 2003.
- [2] Sharif M. et al. Face Recognition: A Survey // Journal of Engineering Science & Technology Review. – 2017. – Т. 10. – №. 2.
- [3] Небаба С.Г. Тестирование технологии подготовки изображений лиц к распознаванию личности в видеопотоке в режиме реального времени // Современная наука: актуальные проблемы теории и практики. Серия «Естественные и технические науки», 2017. №3-4. С. 73-77.
- [4] Небаба С.Г. Захарова А.А. Алгоритм построения деформируемых 3D моделей человеческого лица и обоснование его применимости в системах распознавания личности // Труды СПИИРАН. 2017. Вып. 52. С. 157-179.
- [5] Charles E. Jacobs A. Finkelstein D.H. Sales in Fast Multiresolution Image Querying // Department of Computer Science and Engineering, University of Washington, Seattle, Washington 98195.
- [6] Альтман Е.А., Захаренко Е.И. Быстрый алгоритм вычисления двумерной корреляции для видеообработки // Доклады ТУСУР. 2015. №2 (36) С.119-124.
- [7] Bay H., Ess A., Tuytelaars T., Van Gool L. SURF: Speeded up robust features // Computer Vision and Image Understanding. – 2008. – V. 110. – P. 346–359.
- [8] Leutenegger S. BRISK: Binary Robust invariant scalable keypoints / S. Leutenegger, M.Chli, R.Y. Siegwart // Proceedings of the 2011 International Conference on Computer Vision (ICCV '11), 6 November 2011, P. 2548-2555.
- [9] Фисенко В.Т. Компьютерная обработка и распознавание изображений: учебное пособие / В.Т. Фисенко, Т.Ю. Фисенко - СПб: СПбГУ ИТМО, 2008. – 192 с.
- [10] Монич Ю.И. Оценки качества для анализа цифровых изображений / Ю. И. Монич, В. В. Старовойтов. - Минск: Государственное научное учреждение ОИПИ НАН Беларуси, 2008.
- [11] Kanjar D. Image Sharpness Measure for Blurred Images in Frequency Domain / D. Kanjar, V. Masilamani // International Conference on Design and Manufacturing. - Procedia Engineering, 2013. - P. 149 – 158.
- [12] Tan X. Enhanced Local Texture Feature Sets for Face Recognition Under Difficult Lighting Conditions / X. Tan, B. Triggs // IEEE Transactions on image processing, Vol. 19, № 6, June 2010. – P. 1635-1650.
- [13] Гонсалес Р. Цифровая обработка изображений / Р. Гонсалес, Р. Вудс // Москва: Техносфера, 2012. – 1104 с.
- [14] Фисенко Т.Ю. Исследование и разработка методов улучшения подводных изображений / Т. Ю. Фисенко, В. Т. Фисенко – Материалы международной конференции «Прикладная оптика», 15-19 октября 2012. - С. 294-298.

Об авторах

Небаба Степан Геннадьевич, к.т.н., инженер научно-образовательной лаборатории 3D моделирования и промышленного дизайна Инженерной школы информационных технологий и робототехники Томского политехнического университета. Его e-mail stepanlfx@tpu.ru.