

АЛГОРИТМ СТАБИЛИЗАЦИИ ВИДЕОПОТОКА В ЗАДАЧАХ СПЕЦИАЛИЗАЦИИ «КОМПЬЮТЕРНОЕ ЗРЕНИЕ В АВТОМАТИЗАЦИИ ТЕХНОЛОГИЧЕСКИХ ПРОЦЕССОВ И ПРОИЗВОДСТВ»

В.В. Щербаков, И.В. Шакиров, В.В. Титков
Томский университет систем управления и радиоэлектроники

Рассматривается алгоритм стабилизации видеопотока как методическая разработка для специализации «Компьютерное зрение в автоматизации технологических процессов и производств» в Томском государственном университете систем управления и радиоэлектроники. Основное внимание уделено цифровой обработке изображения, используя перспективное дискретное вейвлет-преобразование.

Ключевые слова: компьютерное зрение, методическая разработка, цифровая обработка изображения.

THE ALGORITHM OF VIDEOSTREAM STABILIZATION IN TASKS OF SPECIALIZATION «COMPUTER VISION IN AUTOMATIONS OF THE TECHNOLOGICAL PROCESSES AND PRODUCTIONS»

V.V. Scherbakov, I.V. Shakirov, V.V. Titkov
Tomsk State University of Control Systems and Radioelectronics

The article discusses the algorithm of video stabilization, as methodological development for specialization «Computer vision in the automation of the technological processes and productions» at Tomsk State University of Control Systems and Electronics. Main attention is paid to digital image processing using promising discrete wavelet transformation.

Keywords: computer vision, methodical development, digital image processing.

Компьютерное зрение – относительно молодая техническая область, которая развивается с 60-х гг. прошлого века. Эта научная дисциплина тесно связана с такими научными областями, как оптика, физика твердого тела, квантовая механика, нейробиология (биологическое зрение), обработка сигналов, статистика, оптимизационная математика, геометрия. Область применения компьютерного зрения довольно широка: медицина (микроскопия, рентгенография, ангиография, ультразвуковые исследования и томография), промышленность (контроль качества, измерение положения и ориентации деталей), военное применение (управление, обнаружение и идентификация объектов), новые области применения – это автономные транспортные средства обнаружения лесных пожаров, нефтяных пятен и др., поддержка создания видеоэффектов для кино и телевидения [1].

Многие проблемы в обучении студентов широкого круга технических специальностей могут быть успешно преодолены с использованием набора новейших программно-алгоритмических средств [2].

Проиллюстрируем это на примере подготовки специалистов по направлению «Компьютерное зрение в автоматизации технологических процессов и производств», которое осуществляется на кафедре электронных средств автоматизации и управления Томского университета систем управления и радиоэлектроники с 2008 г.

На протяжении обучающего процесса был разработан методический материал для изучения задач и проблем компьютерного зрения, а также смежной области – обработки изображений.

Выделены следующие направления подготовки:

- пространственные и частотные методы обработки изображений;
- основы формирования регистрируемых изображений;
- теория распознавания объектов изображения;
- моделирование систем компьютерного зрения;
- методы и алгоритмы сегментации изображений;
- морфологический анализ изображений.

В образовательном процессе очень важно, чтобы студент научился пользоваться и применять комплексно все полученные знания для решения реальных задач.

Одной из таких задач является изучение алгоритма стабилизации видеопотока, изучение решения которой позволяет студентам увидеть связи компьютерного зрения с другими научными смежными областями (рис. 1).

Для решения такой задачи используются статистическая обработка векторов (анализ и получение информации из сигналов, основываясь на их статистических свойствах), цифровая обработка сигналов (преобразование сигналов, представленных в цифровой форме), цифровая обработка изображений. В этом направлении обучаемые более тесно познакомятся с алгоритмами подавления шумов, восстановления изображения и др. Также в обработке сигналов используется дискретное вейвлет-преобразование. Вейвлет-анализ является перспективным направлением цифровой обработки сигналов, его инструменты находят применение в самых различных сферах интеллектуальной деятельности.

Рассмотрим базовую часть задачи стабилизации видеопотока [4]. Одно из основных отличий от профессиональной съёмки от любительской заключается в том, что профессионалы, как правило, снимают со штативами, кранами, тележками и прочими приспособлениями, обеспечивающими неподвижность или плавное движение камеры, тогда как любители обычно об этом не задумываются. Как следствие, часто в любительской съёмке происходит дрожание кадра, а поскольку человек лучше всего реагирует на движение, то хаотичное движение является сильным раздражителем для глаза. Поэтому просмотр клипов с дрожанием очень неприятен для человека.

Видео с дрожанием – это обычно домашнее (любительское) видео, которое появляется при съёмках на фотоаппарат, сотовый телефон или любительскую камеру. На рис. 2 приведена последовательность кадров из ролика с дрожанием.

Введем некоторые обозначения:

- 1) $I_1(x, y), I_2(x, y), \dots, I_i(x, y)$ – кадры видео, где значение в точке определяет ее цвет;
- 2) $T_i = (\Delta x_i, \Delta y_i, s_i, \varphi_i)$ – сдвиг вдоль Ox ,



Рис. 1. Связь между компьютерным зрением и другими областями [3]

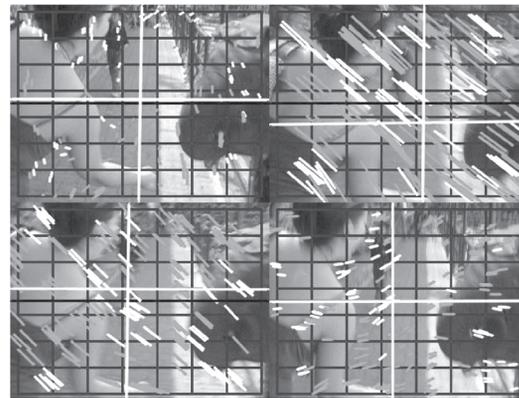


Рис. 2. Последовательность кадров из видео с дрожанием кадра

вдоль Oy , увеличение и поворот i -го кадра относительно $(i-1)$ -го;

3) $\eta_i^* = (\eta_{1i}^*, \eta_{2i}^*, \dots, \eta_{Ni}^*)$ – сдвиг кадра, связанный с дрожанием;

4) $T_i^* = (\Delta x_i^*, \Delta y_i^*, s_i^*, \varphi_i^*)$ – преднамеренный сдвиг кадра (без дрожания).

Значение T_i^* определяется по формуле

$$T_i^* = T_i - \eta_i^*.$$

В данных обозначениях задача формулируется так: необходимо найти T_i^* для всех $i \in [2, N]$ и привести сдвиги между кадрами в соответствие с T^* .

Для завершения картины определим математически понятие сдвига между двумя кадрами [4]:

- 1) пусть $I(x, y), I'(x, y)$ – два кадра;
- 2) $L(\Delta x, \Delta y, s, \varphi)$ – оператор, осуществляющий сдвиг, поворот и увеличение;
- 3) тогда $(\Delta x, \Delta y, s, \varphi)$ определяется по формуле

$$y_{high}[n] = \sum_{k=-\infty}^{\infty} x[k]h[2n - k].$$

Первым и одним из самых сложных и трудозатратных шагов алгоритма является выделение опорных точек на изображении. Для этого шага применяется дискретное вейвлет-преобразование (ДВП), которое нам поможет выделить наиболее значимые детали на изображении.

Определим один уровень ДВП. ДВП сигнала x получаем применением набора фильтров. Сначала сигнал пропускается через низкочастотный (low-pass) фильтр с импульсным откликом g , и получается свёртка:

$$y[n] = (x * g)[n] = \sum_{k=-\infty}^{\infty} x[k]g[n - k].$$

Одновременно сигнал раскладывается с помощью высокочастотного (high-pass) фильтра h . В результате получают детализирующие коэффициенты (после ВЧ-фильтра) и коэффициенты аппроксимации (после НЧ-фильтра). Эти два фильтра связаны между собой и называются квадратурными зеркальными фильтрами (QMF).

Так как половина частотного диапазона сигнала была отфильтрована, то, согласно теореме Котельникова, отсчёты сигналов можно проредить в 2 раза:

$$y_{low}[n] = \sum_{k=-\infty}^{\infty} x[k]g[2n - k],$$

$$y_{high}[n] = \sum_{k=-\infty}^{\infty} x[k]h[2n - k].$$

Такое разложение вдвое уменьшило разрешение по времени в силу прореживания сигнала. Однако каждый из получившихся сигналов представляет половину частотной полосы исходного сигнала, так что частотное разрешение удвоилось.

Таким образом, разложение можно повторять несколько раз для дальнейшего увеличения частотного разрешения с дальнейшим прореживанием коэффициентов после НЧ- и

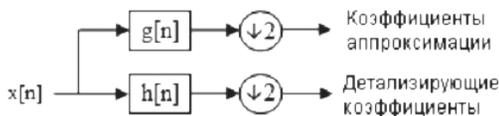


Рис. 3. Схема разложения сигнала в ДВП

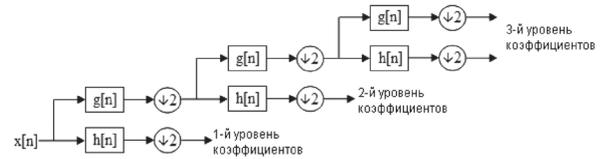


Рис. 4. Дерево разложения

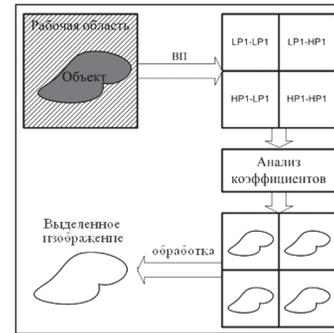


Рис. 5. Пример применения вейвлет-обработки для анализа изображений

ВЧ-фильтрации. Это можно представить в виде двоичного дерева (рис. 4), где листья и узлы соответствуют пространствам с различной частотно-временной локализацией.

На практике алгоритм применения ДВП для изображения можно представить в таком виде, как на рис. 5.

Для поиска опорных точек в алгоритме используется 2-й и 3-й уровни ДВП. Таким образом, избавляемся от огромного количества лишней информации (шум, мелкие детали и т.д.).

Как видно из рис. 2, изображение разделяется на квадранты. В каждом квадранте из массива найденных опорных точек для стабилизации будет использована только одна точка.

Проанализировав два соседних кадра, имеем набор опорных точек. По опорным точкам, изменившим свои координаты, строятся векторы движения. Исходя из найденных векторов, рассчитываются для каждого кадра по одному вектору смещения на каждую ось координат.

Разработанный алгоритм стабилизации видеопотока наглядно демонстрирует применение сразу нескольких методов цифровой обработки изображений, использования вейвлет-анализа, статистической обработки данных (векторов). Также параллельно студенты осваивают несколько основных направлений подготовки по специальности: пространственные и частотные

методы обработки изображений, теория распознавания объектов изображения, методы и алгоритмы сегментации изображений. Кроме того, использование базовых алгоритмов позволяет повысить фундаментальность образования специалистов.

Набор заданий в учебном процессе с использованием таких программно-алгоритмических средств может быть широко использован, а составляющие части алгоритма могут быть задействованы в других смежных дисциплинах, читаемых студентам по другим специальностям. Таким образом, в перспективе студенты будут способны получить новые как общие, так и конкретные знания и применять эти навыки на практике.

Естественным продолжением работы является создание общедоступной и интерактивной версии учебно-методического комплекса, показывающего работу алгоритма стабилизации видеопотока с наглядной демонстрацией работы всех используемых алгоритмов. Это позволит обеспечить рост эффективности обучения. Применение интерактивной версии может

быть полезно как для очной, так и для удаленной подготовки специалистов и предоставит широкие возможности для самообразования студентов, следовательно, позволит повысить информационное наполнение и эффективность использования часа аудиторных занятий. Таким образом, авторы предполагают, что создание подобных общедоступных и интерактивных версий программно-алгоритмических средств позволит перейти на качественно новый уровень обучения.

ЛИТЕРАТУРА

1. *Обзор базовой литературы по компьютерному зрению* / В.В. Титков, В.В. Щербаков, И.В. Шакиров // *Современные техника и технологии*. – Томск, 2009. – Т.1. – С. 237–239.
2. *Арамеев Д.В., Гужов В.И.* Виртуальный доступ к обучающим средам // *Открытое и дистанционное образование*. – Томск, 2008. – № 4 (32). – С. 34–36.
3. http://en.wikipedia.org/wiki/Computer_vision
4. *Стрельников К.* Методы подавления дрожания кадра в видео // *Компьютерная графика и мультимедиа: сетевой журнал [Электронный ресурс]* / *Лаборатория компьютерной графики и мультимедиа, Лаборатории математических методов обработки изображений при ф-те ВМиК МГУ*, 2006. – Режим доступа: <http://cgm.computergraphics.ru/content/view/108>, свободный.