

ПРОБЛЕМЫ РАСПОЗНОВАНИЯ ОБЪЕКТА В ВИДЕОПОТОКЕ

Савич О.А.

Белорусский Национальный Технический Университет, Минск

Компьютерное зрение — теория и технология создания машин, которые могут производить обнаружение, слежение и классификацию объектов в видеопотоке.

Как научная дисциплина, компьютерное зрение относится к теории и технологии создания искусственных систем, которые получают информацию из изображений. Видеоданные могут быть представлены множеством форм, таких как видеопоследовательность, изображения с различных камер или трехмерными данными, например с устройства Kinect или медицинского сканера.

Как технологическая дисциплина, компьютерное зрение стремится применить теории и модели компьютерного зрения к созданию систем компьютерного зрения. Примерами применения таких систем могут быть:

- Системы управления процессами (промышленные роботы, автономные транспортные средства);
- Системы видеонаблюдения;
- Системы организации информации (например, для индексации баз данных изображений);
- Системы моделирования объектов или окружающей среды (анализ медицинских изображений, топографическое моделирование);
- Системы взаимодействия (например, устройства ввода для системы человеко-машинного взаимодействия);
- Системы дополненной реальности;
- Вычислительная фотография, например для мобильных устройств с камерами.

Подразделы компьютерного зрения включают воспроизведение действий, обнаружение событий, слежение, распознавание образов, восстановление изображений и некоторые другие.

Область компьютерного зрения может быть охарактеризована как молодая, разнообразная и динамично развивающаяся. И хотя существуют более ранние работы, можно сказать, что только с конца 1970-х началось интенсивное изучение этой проблемы, когда компьютеры смогли управлять обработкой больших наборов данных, таких как изображения. Однако, эти исследования обычно начинались с других областей, и, следовательно, нет стандартной формулировки проблемы компьютерного зрения. Также, и это даже более важно, нет стандартной формулировки того, как должна решаться проблема компьютерного зрения. Вместо этого, существует масса методов для решения различных строго определённых задач компьютерного зрения, где методы часто зависят от задач и редко могут быть обобщены для широкого круга применения. Многие из методов и приложений все ещё находятся в стадии фундаментальных исследований, но всё большее число методов находит применение в коммерческих продуктах, где они часто составляют часть большей системы, которая может решать сложные задачи (например, в области медицинских изображений или измерения и контроля качества в процессах изготовления). В большинстве практических применений компьютерного зрения компьютеры предварительно запрограммированы для решения отдельных задач, но методы, основанные на знаниях, становятся всё более общими.

Важную часть в области искусственного интеллекта занимает автоматическое планирование или принятие решений в системах, которые могут выполнять механические действия, такие как перемещение робота через некоторую среду. Этот тип обработки обычно нуждается в входных данных, предоставляемых системами компьютерного зрения, действующими как видеосенсор и предоставляющими высокоуровневую информацию о среде и роботе. Другие области, которые иногда описываются как принадлежащие к искусственному интеллекту и которые используются относительно компьютерного зрения, это распознавание образов и обучающие методы. В результате, компьютерное зрение иногда рассматривается как часть области искусственного интеллекта или области компьютерных наук вообще.

Компьютерная графика разделяется на три основных направления: визуализация, обработка изображений и распознавание образов.

Визуализация - это создание изображения на основе некоего описания (модели). К примеру, это может быть отображение графика, схемы, имитация трехмерной виртуальной реальности в компьютерных играх, в системах архитектурного проектирования и т.п.

Основная задача распознавания образов - получение семантического описания изображенных объектов. Цели распознавания могут быть разными: как выделение отдельных элементов на изображении, так и классификация изображения в целом. В какой-то степени задача распознавания является обратной по отношению к задаче визуализации. Области применения - системы распознавания текстов, создание трехмерных моделей человека по фотографиям и т.п.

Обработка изображений отвечает за преобразование (фильтрацию) изображений. Примерами могут служить повышение контраста, резкости, коррекция цветов, сглаживание. Задачей обработки изображения может быть как улучшение (восстановление, реставрация) изображения по какому-то определенному критерию, так и специальное преобразование, кардинально меняющее изображение. В последнем случае обработка изображений может быть промежуточным этапом для дальнейшего распознавания изображения (например, для выделения контура объекта).

Методы обработки изображения могут существенно различаться в зависимости от того, каким путем изображение было получено - синтезировано системой машинной графики, либо путем оцифровки черно-белой или цветной фотографии или видео.

В том случае, если изображение или видеопоследовательность были получены с помощью оцифровки, на них, как правило, присутствует шум. Проблема шумоподавления является одной из самых актуальных и распространенных проблем в области обработки как статичных изображений, так и видео.

Зачем нужно подавлять шум?

Чаще всего шумоподавление служит для улучшения визуального восприятия, но может также использоваться для каких-то специализированных целей - например, в медицине для увеличения четкости изображения на рентгеновских снимках, в качестве предобработки для последующего распознавания и т.п.

Также шумоподавление играет важную роль при сжатии видеопоследовательностей и изображений. И в видео, и в изображениях сжатие основано на пространственной корреляции значений пикселей, а в случае видео - еще и временной корреляции (на сходстве последовательно идущих кадров между собой). Одной из основных проблем в алгоритмах сжатия является определение локальной зашумленности данной области изображения, поскольку при сжатии сильный шум может быть принят за детали изображения, и это может, во-первых, привести к увеличению сложности с точки зрения сжатия и, во-вторых, отрицательно повлиять на результирующее качество сжатого изображения. При сжатии видеоизображений

наличие шума приводит, кроме того, к увеличению межкадровой разницы, понижая, таким образом, степень сжатия, и отрицательно влияет на точность работы алгоритма компенсации движения.

Откуда же берется шум? Источники шума могут быть различными:

- неидеальное оборудование для захвата изображения - видеокамера, сканер и т.п.
- плохие условия съемки - например, сильные шумы, возникающие при ночной фото/видеосъемке;
- помехи при передаче по аналоговым каналам - наводки от источников электромагнитных полей, собственные шумы активных компонентов (усилителей) линии передачи (пример - телевизионный сигнал);
- неточности (плохие фильтры) при выделении яркостного и цветоразностных сигналов из аналогового композитного сигнала и т. п.

Соответственно, шумы тоже бывают разных видов. Самые распространенные:

- белый шум - сигнал, отсчеты которого не коррелируют друг с другом, и его разновидность - белый гауссовский шум, который возникает, в частности, при плохих условиях приема;
- импульсный шум - случайные изолированные точки на изображении, значение которых значительно отличается от значений окружающих их точек (обычно возникает при передаче по аналоговым каналам);
- цветные пятна - характерны для аналогового сигнала (к примеру, присутствуют в видеоизображении, оцифрованном с видеокассет VHS).

Менее распространенные виды шума - шум вида «соль и перец» (как правило, вызывается помехами в электросети) и вертикальные царапины (характерны для старых черно-белых кинолент, возникают при механическом повреждении эмульсии на пленке).

Алгоритмы шумоподавления обычно специализируются на подавлении какого-то конкретного вида шума. Не существует пока универсальных фильтров, детектирующих и подавляющих все виды шумов. Однако многие шумы можно довольно хорошо приблизить моделью белого гауссовского шума, поэтому большинство алгоритмов ориентировано на подавление именно этого вида шума.

Шум в видеопоследовательности можно условно разделить на пространственный и временной. Временным шумом называют неприятный визуальный эффект, возникающий в видео из-за случайного или коррелированного изменения пространственного шума от кадра к кадру. В некоторых источниках под временным шумом понимают любое случайное изменение значений каждого пиксела во времени, в том числе смену яркости от кадра к кадру. Соответственно, фильтры шумоподавления также делятся на пространственные и временные: первые производят усреднение пикселей по каждому отдельному кадру (изображению), а вторые - между несколькими последовательно идущими кадрами. Во временных фильтрах часто используются алгоритмы детектирования и/или компенсации движения. На практике обычно используется сочетание пространственного и временного методов шумоподавления - так называемый, 3D-фильтр. При этом в большинстве существующих алгоритмов первым применяется пространственный фильтр, поскольку после его работы процедура обнаружения движения во временном фильтре становится проще и точнее, и, как следствие, временное шумоподавление становится более эффективным. Хотя обратный порядок применения тоже имеет свои преимущества: временной фильтр, как правило, работает довольно бережно; проработав вначале, он понизит общий уровень шума и предотвратит излишнее размытие, характерное для любого пространственного фильтра.

Основная проблема при пространственном шумоподавлении заключается в том, чтобы не испортить четкость краев предметов на изображении, а также мелкие детали, соизмеримые по амплитуде с шумом. При шумоподавлении в видео подобные детали можно детектировать, отслеживая их на протяжении нескольких кадров.

Основной проблемой при временном шумоподавлении является эффект размытия движения (<ореол>), возникающий при попытке усреднить пиксели вдоль границ движущихся объектов. Для устранения этой проблемы во временные методы часто встраивают алгоритм компенсации движения, но, во-первых, при этом значительно снижается скорость работы фильтра, а во-вторых, неточное нахождение блоков приводит к появлению артефактов.

Еще одной сложностью является оценка качества подавления шума. Как правило, оно оценивается следующим образом: на незашумленное изображение накладывается искусственный шум, затем полученное изображение фильтруется алгоритмом шумоподавления и сравнивается с исходным с помощью какой-нибудь метрики. Чаще всего для этой цели используют метрику PSNR (peak signal-to-noise ratio - пиковое соотношение сигнал/шум).

Однако эта и другие аналогичные метрики позволяют оценивать лишь среднеквадратичную разницу между изображениями, поэтому лучшие результаты с точки зрения метрики не всегда соответствуют наилучшему визуальному восприятию. К примеру, более размытое изображение, в котором вместе с шумом были удалены мелкие детали, может иметь значение PSNR выше, чем изображение с более аккуратно подавленным шумом. Визуальные оценки при этом покажут предпочтительность второго изображения.

При оценке качества работы временного алгоритма шумоподавления часто рассматривается межкадровая разница - чем меньше она становится после обработки видео фильтром шумоподавления, тем лучше считается качество этого фильтра.

Различные существующие на данный момент подходы к шумоподавлению будут рассмотрены далее.

Можно выделить следующие базовые подходы к пространственному шумоподавлению:

- Линейное усреднение пикселей по соседям;
- Медианная фильтрация;
- Математическая морфология;
- Гауссовское размытие;
- Методы на основе вейвлет-преобразования;
- Метод главных компонент;
- Анизотропная диффузия;
- Фильтры Винера.

Некоторые из этих методов применимы с небольшими модификациями также и во временной области.

Заметим, что алгоритмы на основе вейвлет-преобразования и метода главных компонент применяются, в основном, для обработки статичных изображений, хотя и обеспечивают наилучшее качество среди всех вышеперечисленных методов. Дело в том, что эти алгоритмы работают очень медленно и даже при хорошей оптимизации не могут обеспечить обработку в реальном времени, а при обработке видео скорость играет очень важную роль.

ЛИТЕРАТУРА

1.Szymon Graboskwi, Wojciech Bienieck "A two-pass median-like filter for impulse noise removal in multi-channel images". *KOSYR 2003, str. 195-200.*

2. Detlev Marpe, Hans L. Cycon, Gunther Zander, Kai-Uwe Barthel "Context-based Denoising of Images Using Iterative Wavelet Thresholding". *Proc. SPIE, Vol. 4671, pp. 907-914, Jan. 2002.*
3. D. Darian Muresan, Thomas W. Parks "Adaptive principal components and image denoising". *IEEE ICIP 2003.*
4. Ivan W. Selesnick and Ke Yong Li "Video denoising using 2D and 3D dual-tree complex wavelet transforms". *Appl Signal Image Proc. X (Proc. SPIE 5207), Aug 2003.*