

В.А. Минаев
ПРОЕКТ VSAM – НОВЫЙ ПОДХОД
К ОБЕСПЕЧЕНИЮ БЕЗОПАСНОСТИ

Проект VSAM (Video Surveillance and Monitoring) выполнен Институтом робототехники Университета Карнеги-Меллона и Корпорацией Сарнофф. Его суть заключалась в реализации системы видеонаблюдения и мониторинга, использующей множество интеллектуальных видеодатчиков для контроля людей и транспортных средств в условиях сложной окружающей среды.

В результате исследования апробирована технология, позволяющая единственному человеку-оператору контролировать ситуацию на протяженных объектах, используя распределенную сеть датчиков и интеллектуальный комплекс обработки данных, поступающих от них [1].

Эта технология может быть эффективно применена для контроля сложных периметров, обработки данных с беспилотных самолетов-разведчиков, обеспечения безопасности посольств и аэропортов, выявления подозреваемых в распространении наркотиков или возможных террористов путем сбора и обработки видеоданных, собранных в определенных зонах. Нужно также отметить, что коммерческий сектор также является важной областью использования данной технологии.

Необходимость в ней назрела и в связи с тем, что установка камер для получения видеообразов становится все дешевле, а использование людей для визуального анализа в реальном масштабе времени – все дороже.

Сейчас камеры видеонаблюдения стоят практически во всех уважающих себя учреждениях, пишутся десятки часов видеоархивов. Однако реакция на факты правонарушений и других инцидентов происходит лишь после того, как все произошло – склад ограблен, автомобиль украден. И, как правило, реакция бывает запоздалой. Технология же VSAM продвигает нас к тому времени, когда анализ видеоданных позволит предупредить преступление.

Используя технологию VSAM, можно автоматически классифицировать людей и транспортные средства, определять их местоположение, выделять движущиеся объекты. Все обнаруженные системой цели сначала классифицируются в семантические категории типа “человек”, “группа людей”, “автомобиль”, используя форму и цветовой анализ. В дальнейшем происходит, например, применительно к людям, классификация человеческой деятельности по типу “ходьба”, “бег”, “прыжки”.

В конечном итоге эти семантические категории отображаются в сим-

волическом виде на графической карте, удобной для восприятия человека-оператора. Система VSAM отлаживалась на территории городка университета Карнеги-Меллона. Ее общая структура показана на рис. 1.

Интеллектуальные сенсоры (ИС) выступают своеобразным интеллектуальным фильтром между видеокамерами и VSAM-сетью. Его функция – выделение существенных (в зависимости от постановки задачи) объектов или событий в видео образах и передача этой информации в символьном виде в центральную систему обработки (ЦСО). Такой подход резко уменьшает требования к полосе пропускания VSAM-сети.

Во VSAM-системе могут использоваться самые разнообразные интеллектуальные датчики и периферийные устройства. Например, цветные камеры с зарядовой связью, тепловые датчики, устройства фокусировки и масштабирования, ИК-камеры. Для измерения местоположения применяется глобальная система позиционирования с использованием спутников (GPS).

Работа ЦСО связана с получением обработанных видеоданных от каждого ИС, объединением их с модельной информацией о наблюдаемом участке и информацией из базы данных об известных объектах и выбором ситуаций, представляющих интерес для пользователя. Соответствующие результаты передаются на компьютеризированное рабочее место с графическим интерфейсом или на любое другое средство визуализации.

Структура сети VSAM такова, что позволяет нескольким ЦСО обмениваться информацией между собой. Для этого разработан специальный коммуникационный протокол.

Ключевым элементом ЦСО выступает система контроля ИС. Ее основное назначение – гарантированное включение ресурсов системы наружного видеонаблюдения с целью качественного решения всех задач, оставленных пользователем, и слежение за всеми важными объектами в любой момент времени. Эти задачи ставятся пользователем, используя возможности, предоставляемые средствами графического интерфейса, и включают конкретные объекты (цели), определенные зоны наблюдения, значимые события (типа человека, подозрительно слоняющегося около двери в спецпомещение). Система контроля ИС определяет оптимальную комбинацию последних с целью максимизации выполнения требований к работе VSAM.

Очевидно, что мониторинг передвижения множества людей и транспортных средств в сложной городской среде – невыполнимая задача для одного человека-оператора, когда используется обычная технология – наблюдение за десятками экранов, показывающих “живую” видеоинформа-

цию. Такой подход гарантирует только одно – потерю значимой, а подчас – ключевой информации, а также необходимость в использовании широкополосных (а значит – относительно дорогостоящих) каналов ее передачи.

Идея VSAM-подхода состоит в том, чтобы обеспечить диалоговый, графический пользовательский интерфейс, при котором реальные объекты (люди, машины и т.п.) автоматически замещаются на их образы, “живущие” в синтезированном мире окружающей среды.

Этот подход имеет огромные преимущества, связанные с тем, что визуализация событий не привязана к оригинальной, переусложненной событиями, людьми и другими объектами, картине и “видению” единственного видеодатчика.

Средства графического интерфейса позволяют управлять задачами как конкретных интеллектуальных сенсоров, так и их групп. Это дает возможность пользователю охватывать всю постоянно изменяющуюся картину: объекты, сенсоры и зоны наблюдения. При этом с помощью так называемых “пусковых механизмов” изменяются сектора обзоров видеоперенос сенсоров, производится масштабирование, привязка сенсоров к наблюдению за конкретными движущимися объектами и т.п.

В технологии VSAM применены специальные робастные алгоритмы для обнаружения движущихся объектов и слежения за ними.

Известно, что задачу обнаружения движения можно выполнить с использованием трех подходов: временного вычитания кадров [2], вычитания бэкграунда [3, 4] и нахождения оптического потока [5]. Каждый из этих методов имеет и достоинства, и недостатки.

Метод вычитания кадров очень адаптивен к динамическим изменениям среды, но мало пригоден для выявления релевантной информации об особенностях изображения.

Метод вычитания бэкграунда обеспечивает наиболее полное понимание особенностей изображения, однако он чрезвычайно чувствителен к изменениям изображения, обусловленным колебаниями освещенности и другими подобными событиями.

Метод оптического потока может использоваться, чтобы обнаружитьдвигающиеся объекты при движении камеры, но в вычислительном отношении он сложен и не может быть применен к полным видеопотокам в реальном времени без специализированных ЭВМ.

Исходя из этого, в рамках проекта VSAM разработан и применен новый алгоритм обнаружения движения: комбинация адаптивного вычитания бэкграунда и вычитания трех кадров – этот гибридный алгоритм оказался

быстрым и удивительно эффективным [6].

В заключение хотелось бы отметить, что в результате реализации проекта VSAM выявлены направления развития “нового видео”, основанного на революционной интеллектуализации его возможностей. Будущая работа исследователей будет связана с:

- лучшим пониманием особенностей человеческого движения, включая сегментацию и трекинг частей тела;
- улучшение геомodelей;
- совершенствованием процедур обнаружения и классификации многообъектных событий и действий;
- совершенствованием управления камерами.

Литература

1. T. Kanade, R. Collins, A. Lipton, P. Anandan, and P. Burt. Cooperative multisensor video surveillance. In Proceedings of the 1997 DARPA Image Understanding Workshop, volume 1, pages 3–10, May 1997.
2. C. Anderson, Peter Burt, and G. van der Wal. Change detection and tracking using pyramid transformation techniques. In Proceedings of SPIE - Intelligent Robots and Computer Vision, volume 579, pages 72–78, 1985.
3. I. Haritaoglu, Larry S. Davis, and D. Harwood. Who? When? Where? What? A real time system for detecting and tracking people. In FGR98, 1998.
4. C. Wren, A. Azarbayejani, T. Darrell, and Alex Pentland. Real-time tracking of the human body. IEEE Transactions on Pattern Analysis and Machine Intelligence, 19(7):780–785, 1997. Robotics Institute, CMU – 68 – VSAM Final Report.
5. J. Barron, D. Fleet, and S. Beauchemin. Performance of optical flow techniques. International Journal of Computer Vision, 12(1):42–77, 1994.
6. T. Kanade, R. Collins, A. Lipton, P. Burt, and L. Wixson. Advances in cooperative multisensor video surveillance. In Proceedings of the 1998 DARPA Image Understanding Workshop, volume 1, pages 3–24, November 1998.