

Использование web-камер как источника стереоизображений в реальном времени.

Протасов С.И. e-mail: stanislav.protasov@gmail.com

Воронежский государственный университет

В настоящее время формирование и обработка стереоизображений являются очень популярным направлением в отрасли развлечений (фильмы и компьютерные игры в формате 3D, игровые сенсоры, расширенная реальность), образования (виртуальные стенды) и промышленности (компьютерное зрение в технологических процессах). Использование технологий трёхмерного видео (затворное, анаглифическое, поляризованное) позволяет наиболее близким к естественному способом сформировать у человека ощущение реальности виртуальных объектов. Информация, содержащаяся в стереопаре, предоставляет возможность не только создавать эффект трёхмерности объектов при просмотре, но и формировать виртуальные трёхмерные модели, предназначенные для последующей обработки и передачи по каналам связи.

Стоимость откалиброванных камер даже невысокого разрешения, снимающих стереографическое видео является на настоящий момент высокой. Использование пары обычных камер для стереосъёмки сопряжено с рядом проблем. Для того чтобы сформированные двумя камерами изображения можно было считать стереографическими, они должны пройти предварительную обработку. Для получения стереоизображений при помощи пары камер, необходимо пройти следующие этапы:

- калибровка каждой из камер (определение внутренних параметров камеры);
- стерео-калибровка камер (калибровка в общей системе координат);
- ректификация изображений;
- синхронизация камер.

Для решения первых двух задач, связанных с калибровкой, существуют готовые алгоритмические и программные решения (Matlab Camera Calibration Toolbox, библиотека OpenCV). Результатом использования алгоритмов калибровки являются калибровочная матрица [1] и данные о дисторсии. Стерео-калибровка предполагает получение матрицы, описывающей перенос точки из системы координат одной камеры в систему координат другой. **В случае, если камеры неподвижно зафиксированы друг относительно друга и в их настройках отключены опции автоматической фокусировки,**

получаемые в ходе калибровки данные могут быть вычислены однократно и представлены в удобном для обработки виде. Так, с помощью матрицы внешней калибровки мы предварительно вычисляем соответствующие эпиполярные линии на изображениях для последующего быстрого выполнения ректификации.

Под ректификацией обычно понимают проецирование двух или более изображений на одну плоскость [2] обычно так, чтобы строки изображений соответствовали эпиполярным линиям [3]. Дополнительное требование вводится для того, чтобы изображения легче обрабатывались алгоритмами стерео-сопоставления. Программная реализация ректификации может быть хорошо распараллелена на произвольное количество независимых потоков, что позволяет максимально использовать тенденции в области архитектуры современных центральных и графических процессоров.

Задача программной синхронизации web-камер является нетривиальной. С точки зрения качества получаемого результата очевидным кажется использование идентичных камер, однако хотя это исключает возможность использования одной USB-шины. Нами предложена модель синхронизации идентичных web-камер, использующих независимые USB-шины, позволяющая получать синхронизированные стереоизображения с высокой частотой.

Полностью избавиться от рассинхронизации камер не представляется возможным, это связано с многозадачностью современных операционных систем и функционированием планировщика заданий, однако её можно минимизировать. Экспертная оценка допускает величину рассинхронизации камер не выше 10 мс.

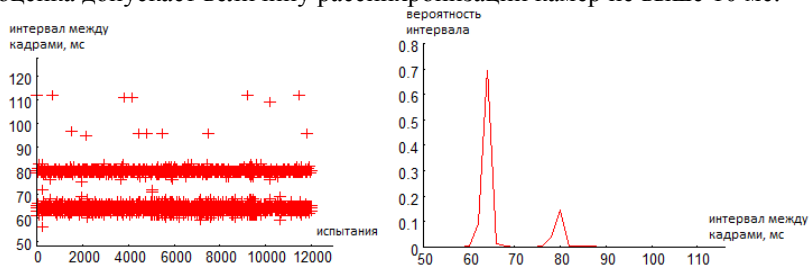


Рисунок 1. Выборка интервалов между последовательными кадрами и распределение их вероятностей

На рис. 1 представлена выборка интервалов между последовательным захватом кадров одной камерой. Величины интервалов кластеризуются вокруг значений, кратных некоторой

величине, в данном случае 16 мс, что связано с устройством web-камер. Разброс значений в кластере обусловлен работой планировщика заданий операционной системы. Попадание интервала в доминирующий кластер вокруг величины 64 мс в нашем случае имеет вероятность 81.5%. Таким образом, можно построить модель временных характеристик камеры, опираясь на следующие значения:

- время инициализации камеры S_1 и S_2 (от начала инициализации до первого кадра);
- интервалы T_1 и T_2 , которым кратны середины кластера;
- доминирующие кластеры интервалов $\{K_1\}_i$ и $\{K_2\}_j$ (обычно их не более двух), и их вероятности $\{P_{K1}\}_i$ и $\{P_{K2}\}_j$, а также разброс значений $\pm\delta$ в кластере.

Получены эмпирические результаты, показывающие, что при искусственном простое в захвате камеры величина простоя также будет кратна T_i .

Предложенная модель позволяет предсказывать время появления очередного кадра и принимать меры по восстановлению синхронности. Мы предлагаем следующий алгоритм для поддержания синхронности захвата изображений камерами:

1. Инициализировать камеры с учётом значений времени инициализации S_1 и S_2 так, чтобы время рассинхронизация захвата первого кадра была минимальной.
2. В случае если камеры имеют различные доминирующие кластеры, предлагается использовать искусственный простой величины $|\{K_1\}_1 - \{K_2\}_1|$ на каждом шаге для более быстрой камеры.
3. В случае запаздывания одной из камер на величину t , превышающую $T_i/2$, инициировать искусственный простой опережающей камера на один (или более) интервал T_j . Таким образом, с вероятностью $(\{P_{K1}\}_1)^2$ рассинхронизация следующих кадров будет иметь величину $(T_i - t) \pm \delta$, не превышающую $T_i/2$.

Таким образом, при использовании идентичных камер можно гарантировать, что рассинхронизация между кадрами с обеих камер не будет превышать $T_i/2 = 8$ мс с вероятностью $\{P_{K1}\}_1 * \{P_{K2}\}_1 = (\{P_{K1}\}_1)^2 = 0.815^2 = 0.66425$. При средней частоте обновления кадра 14.9 fps можно получить до 10 fps синхронизированного видео-потока.

На основании приведённой модели построена экспериментальная система формирования анаглифических изображений в реальном времени. Система состоит из двух камер Logitech C270 и ноутбука на базе процессора Intel Core i7. Программная часть реализована для

платформы .NET 3.5 с использованием параллельных вычислений и unsafe-кода. Эмпирически показано, что вычислительной мощности персонального компьютера достаточно для выполнения ректификации и анаглифического слияния в реальном времени.

Литература

1. Вахитов А.Т. Обзор алгоритмов стереозрения / А.Т. Вахитов, Л.С. Гуревич, Д.В. Павленко. Стохастическая оптимизация в информатике, Вып. 4 / Под ред. О. Н. Граничина - СПб.: Издательство Санкт-Петербургского университета 2008. - 299с.
2. Linda G. Shapiro and George C. Stockman (2001). Computer Vision. Prentice Hall. pp. 580.
3. Стерео-реконструкция / А. Конушин [и др.] // МГУ, ВМиК. – 2008. – (<http://courses.graphicon.ru/main/vision2008>)