

Универсальный видеокодек реального времени на основе трехмерного дискретного косинусного преобразования

Ш. С. Фахми¹, С. В. Очкур²

¹СПБГЭТУ «ЛЭТИ» кафедра САПР, ²СПБГЭТУ «ЛЭТИ» кафедра ТВ

В основе современных видеокодеков, реализующих оптимальное согласование источников видеoinформации с каналами связи путем компрессии/декомпрессии видеoinформации, лежит принцип устранения статистической избыточности (декорреляции). Принято различать пространственную (внутрикадровую) и временную (межкадровую) избыточность.

Для устранения пространственной избыточности наряду с другими широко используются алгоритмы на основе дискретного косинусного преобразования (DCT).

Для устранения временной избыточности, как правило, применяются алгоритмы вычисления межкадровой разности и компенсации движения. К числу недостатков последних принято относить их высокую вычислительную сложность, существенно затрудняющую их практическую реализацию. В некоторых предметных областях вычислительная сложность указанных алгоритмов становится непреодолимым препятствием. Наиболее характерным примером являются системы реального времени. При этом обычно умалчивается другой важный недостаток подобных средств – они являются проблемно ориентированными, т. е. не универсальными. Другими словами, они не могут быть эффективно использованы для широкого класса практически интересных задач.

Высокая вычислительная сложность большинства современных алгоритмов компенсации движения является следствием природы этих задач. С теоретической точки зрения эти задачи с большой степенью уверенности можно отнести к классу так называемых «трудно формализуемых» задач, для которых характерна следующая особенность. При формализации любой задачи для уточнения понятия цели явно или неявно должна задаваться некоторая оценочная функция (функция полезности), выражающая степень приближения к цели или устанавливающая порядок предпочтения возможных вариантов достижения цели. Для трудно формализуемых задач полностью определить такую функцию либо невозможно, либо связанные с ней вычисления и алгоритмы чрезвычайно громоздки. Эти особенности в полной мере относятся к средствам временной декорреляции, определяемым стандартами MPEG-2, MPEG-4 (H.264). Полный перебор векторов движения для H.264 требует более 10^{22} операций в секунду (только арифметических, без учета обмена данными с памятью) для видео стандартного разрешения. Поэтому обычно применяют упрощенные алгоритмы поиска, которые позволяют снизить сложность вычислений. Однако затраты на их реализацию все равно остаются значительными и составляют 60% - 96% от всех вычислительных затрат. При этом упрощенные алгоритмы поиска более чем в 3 раза ухудшают эффективность кодирования за счет пропуска так называемых локальных минимумов.

В контексте необходимости декорреляции, видеосигнал как многомерная функция имеет оптимальное разложение по всем своим аргументам. Другими словами, для

устранения временной корреляции нет существенных формальных оснований использовать средства, отличающиеся от применяемых для пространственной декорреляции. Кроме того, известно, что для большинства практически значимых классов изображений оптимальным (асимптотически оптимальным) по критерию минимума требуемого для передачи количества информации при заданном уровне качества изображения является DCT.

Достоинством видеокодека реального времени на основе 3D-DCT наряду с универсальностью (отсутствие потребности вычисления оценок движения) является возможность его параметрической адаптации к интервалам корреляции видеоинформации по различным аргументам без перестройки структуры видеокодека, обеспечивающей наилучшее приближение к предельной скорости передачи информации при заданных ограничениях на качество передачи и сложность кодека. Исследования показали, что при такой адаптации должны изменяться размеры пространственных и временных интервалов вычисления 3D-DCT, которые определяются соответствующими интервалами корреляции. Для большинства практически значимых классов изображений размеры этих интервалов (представленные как целочисленные степени числа 2) составили от 8×8 до 64×64 пикселей и от 4 до 128 кадров.

Реализация и сравнительные испытания 3D-DCT видеокодека показали перспективность этого класса устройств, характеризующегося универсальностью, адаптивностью и высокой производительностью, обеспечивающими наилучшее приближение к предельной скорости передачи информации при заданных ограничениях на качество передачи и сложность кодека. Можно утверждать, что 3D-DCT видеокодеки целесообразно применять в телевизионных системах различных форматов кадра и различных назначений.