

В.Г.ИВАНОВ, канд. техн. наук

ТЕХНОЛОГИИ КОДИРОВАНИЯ СИГНАЛОВ РАЗЛИЧНОЙ ФИЗИЧЕСКОЙ ПРИРОДЫ

У роботі науково обґрунтовується і пропонується технологія створення автоматизованих комп'ютерних систем уведення, кодування, збереження, розпізнавання і пошуку сигналів і зображень різноманітних класів: мовних сигналів і зображення протекторів шин автомобільних коліс та відбитків печаток і штампів.

Современные информационные технологии на базе глобальной компьютерной сети Internet предоставляют пользователям широкие возможности в получении информации самого разного толка. Можно загрузить музыкальные произведения со специальных сайтов, посмотреть мультимедиа презентации, озвученные Web-странички. Разрабатываются и размещаются в сети Internet так называемые интерактивные виртуальные студии, которые позволяют в реальном времени аранжировать несколько звуковых фрагментов в единую музыкальную композицию (<http://www.erack.com/qweb/studio/studio.htm>).

Однако все это требует передачи по сети гигантского количества мультимедийной информации. Так одна минута цифровой аудиозаписи с качеством лазерного диска занимает порядка 10,8 Мб информации. Несложно подсчитать, что для прослушивания в реальном времени такого файла нужно иметь скорость передачи до 1,5 Мб/сек.

Трудно надеяться, что отечественные, даже выделенные телефонные каналы обеспечат скорость передачи больше, чем 16-32 Кб/сек. Поэтому, чтобы сделать этот вид услуг доступным широкому пользователю, многие фирмы стали разрабатывать свои алгоритмы и форматы сжатия изображения и звука.

За несколько лет работы над проблемой компрессии аудиои видеоинформации накопилось множество продуктов, методов, алгоритмов, которые пересекаются между собой, зачастую имеют по несколько разных названий и функций. Большинство из них базируется на сходных принципах, а оригинальных алгоритмов разработано не так уж и много [1, 2]. Можно выделить следующие методы: а) группового кодирования (RLE); б) Лемпела-Зива-Велча (LZW); в) CCITT; г) дискретных косинус-преобразований (DCT), применяемых в JPEG-сжатии; д) JBIG; е) ART; ж) фрактального сжатия.

Групповое кодирование (RLE) – алгоритм сжатия данных, поддерживаемый большинством растровых файловых форматов, таких как

TIFF, BMP и PCX. Алгоритм RLE позволяет сжимать данные любых типов. Однако сама информация влияет на полноту сжатия и большинство алгоритмов не могут достигать высокой степени сжатия.

Схема сжатия LZW является одной из наиболее распространенных в компьютерной графике. Этот метод сжатия данных без потерь базируется на словарях и применяется в различных форматах файлов изображений (GIF, TIFF) и включен в стандарт сжатия для модемов V.42 bis и PostScript Level 2.

LZW позволяет хорошо сжимать данные различной пиксельной глубины, но изображения с шумом могут значительно уменьшить эффективность сжатия при помощи LZW.

Кодирование CCITT - форма сжатия данных, применяемая для факсимильной передачи и стандартизованная Международным консультативным комитетом по телеграфии и телефонии (CCITT). Стандарт базируется на схеме ключевого сжатия, предложенной Дэвидом Хаффменом и широко известен как кодирование по алгоритму Хоффмена.

Алгоритм, разработанный объединенной экспертной группой по фотографии (JPEG) использует схему кодирования на базе дискретных косинус-преобразований (DCT). На сегодняшний день JPEG является одним из наиболее актуальных направлений развития технологии сжатия изображений. Схема JPEG была специально разработана для сжатия цветных и полутоновых многоградационных изображений – фотографий, телевизионных заставок, другой сложной графики. Однако анимация, черно-белые иллюстрации и документы сжимаются плохо.

Фрактальное кодирование основано на том факте, что естественные и большинство искусственных объектов содержат избыточную информацию в виде одинаковых, повторяющихся рисунков, которые называются фракталами. Однако процесс кодирования требует исключительно большого объема вычислений.

Все выше сказанное опять подтверждает важность и актуальность исследований вопросов экономного представления и кодирования сигналов различной физической природы, которые ведутся нами в рамках решения практических задач создания многофункциональных систем обработки информации.

Инструментальное исследование речевых сигналов вообще и в криминалистике в частности является чрезвычайно сложной и далекой от полного решения проблемой, которая включает в себя задачу установления личности человека по фонограмме его устной речи.

Содержанием этих видов исследований является изучение элементов и признаков устной речи, связанных со свойствами человека и позволяющих осуществить идентификацию и диагностику личности говорившего. При идентификации экспертизой решается вопрос, является

ли устная речь на сравниваемых фонограммах устной речью одного и того же человека. В задачи диагностических исследований входит распознавание по фонограмме различных социальных и психологических характеристик личности говорившего, имеющих существенное значение для розыскных и следственных действий.

Современная теория и практика обработки речевых сигналов базируется в основном на спектрально-временном представлении звуков речи, а эффективность распознавания или идентификации основывается на правильном или удачном выборе признаков речевого сигнала, т.е. совокупности параметров, которая непосредственно используется при анализе и принятии решений. Формального подхода для выбора признаков при решении той или иной задачи обработки речевых сигналов на наш взгляд нет. Хотя можно сформулировать ряд общих требований при их выборе. Признаки должны обладать максимальной информативностью, простотой измерения, стабильностью во времени, частотой и естественностью появлений в речи, инвариантностью к возмущающей среде, невосприимчивостью к имитации. Понятно, что найти признаки, отвечающие всем этим требованиям чрезвычайно трудно, если вообще возможно.

В разрабатываемой нами системе в качестве признаков используются и развиваются результаты обработки сигналов различной физической природы с использованием класса обобщенных базисно-ограниченных ортогональных преобразований Фурье-Хаара для решения задач сжатия данных [3, 4].

Анализ и сравнение речевых сигналов (полных слов) дикторов производится по следующим параметрам: прямой спектр сигналов; амплитудно-упорядоченный спектр; прямой коэффициент сжатия сигналов; упорядоченный коэффициент сжатия; отношение числа положительных составляющих спектра Хаара к числу отрицательных составляющих; мощность основного участка спектра. Речевой материал (фонограммы) был представлен на дискетах и представлял собой фонограммы, наговоренные с микрофона различными дикторами в одинаковых условиях. Каждый диктор произносил по два раза подряд один и тот же список слов (цифры от одного до десяти и фразу «сегодня вечером»).

Предварительный анализ полученных результатов позволяет сказать следующее. Графическое отображение амплитудных спектров для трех различных дикторов, произносящих цифры от одного до десяти показывает относительно устойчивую формантную картину. Это характерно как для базиса Фурье, так и для базиса Хаара. Причем для спектра Хаара можно отметить наличие заметных формантных областей в высокочастотной области спектра. На наш взгляд хорошее совпадение

спектральных картин в базисе Хаара дают произносимые диктором слова: «два», «пять», «шесть», т.е. когда говорятся односложные слова или ударение падает на первый слог. Спектральный отпечаток слова «четыре» характеризуется большим количеством неравномерных формантных областей, что затрудняет визуальный анализ на совпадение. Эффективность сжатого представления сигналов в базисе Хаара выше, чем в базисе Фурье, что объясняется нестационарной природой функций Хаара и самих речевых сигналов.

Одной из актуальных задач судопроизводства на современном этапе является повышение роли использования научно-технических средств и методов в раскрытии и расследовании преступлений. Ведущая роль в этой работе принадлежит экспертным учреждениям Министерства юстиции Украины, а также экспертно-криминалистическим подразделениям Министерства внутренних дел. Однако, эта роль в настоящее время не может быть эффективно реализована ввиду отсутствия современных, быстро пополняемых информационно-справочных баз видео данных (рисунок протекторов шин автомобильных колес, конфигурация и рисунок моделей фарного стекла, рисунок подошвы обуви и т.д.). Использование данных на классических носителях, таких, как справочные издания, справочные таблицы и тому подобное, не отвечает современным требованиям ввиду их высокой стоимости и сложности в обновлении и поиске информации.

Поэтому проводимые нами работы по созданию таких баз видеоданных являются чрезвычайно необходимыми и важными исследованиями. Эти работы включают в себя очень широкий круг непростых вопросов сбора необходимой информации, ввода изображений типа портрет и символьной графической информации в память ЭВМ, фильтрации и компактного представления изображений, формирования и выбор признаков с целью автоматизации криминалистической трасологической экспертизы.

Список литературы: 1. *Маркел Дж.Д., Грей А.Х.* Линейное предсказание речи: Пер. с англ. /Под ред. Ю.Н.Прохорова и В.С.Звездина. – М.: Связь, 1980. – 308 с. 2. *Д.Мюррей, У ван Райнер.* Энциклопедия форматов графических файлов: пер. с англ. – К.: Издательская группа ВНУ, 1997. – 672 с. 3. *Иванов В.Г.* Компьютерный анализ ортогональных преобразований в задачах сжатия дискретных сигналов // Компьютерные методы исследования проблем теории и техники передачи сигналов по радиоканалам: Тез. докл. Всесоюз. научн.-техн. конф. Евпатория, сентябрь, 1990. - М.: Радио и связь, 1990. 4. *Иванов В.Г.* Эффективность анализа и синтеза в дискретном базисе Хаара// Радиоэлектроника – 1983. - № 9. – С.54 – 56 (Изв. высш. учеб. заведений).

Поступила в редколлегию 16.05.2000