

Эссе по курсу "Защита Информации" на тему:

**МОДУЛЯЦИЯ ВРЕМЕНИ: НОВЫЙ ПОДХОД В
СОЗДАНИИ ВОДЯНЫХ ЗНАКОВ ДЛЯ АУДИО И
ИЗОБРАЖЕНИЯ**

Выполнил студент 011 группы
КАЛИНЧУК СЕРГЕЙ

МФТИ
г. Долгопрудный
2004 г

Рассматривается новый метод для сокрытия информации в аудио(звукозаписи) и изображении с помощью таких "незаметных" сжатий и растяжений временных областей в аудио файле, что внесенные ими изменения были бы не слышны. При сравнении измененного файла с опорной (reference) копией, то есть с первоначальным файлом до изменений, можно обнаружить области сжатия и растяжения. Местоположение и длительность этих преобразований представляют собой закодированную информацию в модифицированном файле. Этот новый подход имеет много преимуществ над другими методами. Принципиальное отличие от большей части watermarking-методов (методов, которые создают водяные знаки) — это то, что он имеет теоретически пониженный уровень шума, не вводит спектральных искажений и устойчив ко всем схемам сжатия и передачи. Также можно сделать, чтобы в этом методе файл был самосинхронизирующимся, так что он не будет зависеть от опорного файла.

Введение

Существует много способов для изменения масштаба временной оси (time-scale modification) в аудио. Этот метод имеет прямое отношение к изменению длительности аудио образца без изменений основного тона (речевого сигнала) и других спектральных характеристик. Когда основной тон не меняется, небольшие количества модификаций масштаба времени обычно незаметны, что является базисом в подходе создания водяных знаков. Модулируя масштаб времени аудио файла, информация может быть незаметно закодирована в него. Как показано на рисунке 1, короткие временные области файла или сжимаются или растягиваются на незначительное количество (преувеличение на рисунке для иллюстрации). Этот метод называется "time base modulation", когда базовая временная ось модулируется watermark-функцией(создающей водяной знак). При балансировке количества сжатий и растяжений общая длина файла не изменяется. Местоположение и степень сжатия или растяжения являются величинами, которые используются для кодирования информации. Водяной знак обнаруживается сравнением watermarked-копии(файл, который уже закодировали) с опорным аудио. Выравнивание по времени watermarked- и опорного аудио создает "tempo map" (скоростная карта), которая показывает как временной базис watermarked-аудио был изменен. В областях сжатия или растяжения tempo map отклонятся от прямой линии, и вставленные данные водяного знака могут быть восстановлены из этих отклонений. Подобный метод можно использовать в создании водяного знака для изображения или отсканированного текста с помощью тех же "невидимых" сжатий и растяжений областей изображения.

Водяной знак может кодировать информацию об авторском праве, криптографическую подпись или информацию, которая специально распознает определенную индивидуальную копию аудио или изображения. Это весьма полезно, например, чтобы скрыть ключи шифрования или цифровые права на пользование. Если каждому легальному пользователю дан файл, охраняемый законом об авторском праве, с уникальным водяным знаком, то в незаконно распространенных копиях можно найти этот знак и по нему определить источник "утечки". Другое применение заключается в кодировании криптографического hash-сигнала, чтобы проверить его подлинность, то есть своего рода применение в цифровой подписи. Любое изменение сигнала (такое как вставка или удаление данных) будет генерировать другую hash величину, которую можно сравнить с раскодированной величиной для обнаружения вмешательства. Водяные знаки также часто применяют в приложениях с высоким уровнем секретности (индивидуализация выданных документов). В последней части будет рассмотрено, как процедура изменения оси времени может быть встроена прямо в ксерокс или принтер, позволяя каждой отдельной копии быть уникально подписанной. Есть много различных моделей сжатия и растяжения, применяемых в кодировании информации. Бывают случаи, когда желательно иметь одинаковое количество сжатий и растяжений так, чтобы watermarked-версия была такого же размера, как и опорный файл, хотя это

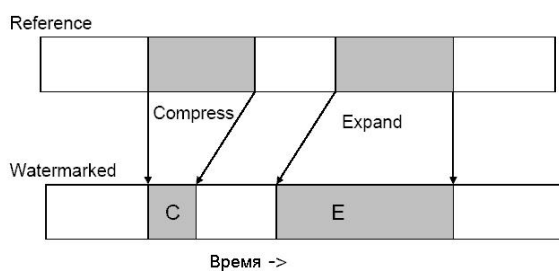


Рисунок 1. 'Watermarking' сигнала при модуляции 'времени'.

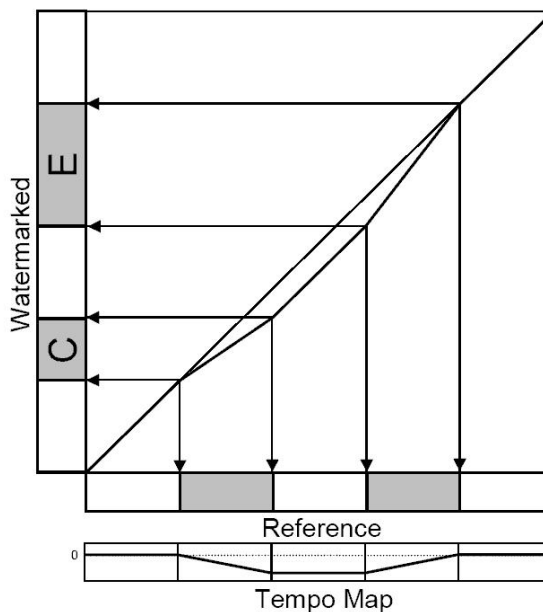


Рисунок 2. Восстановление "скоростной" карты ("tempo map") с помощью динамического выравнивания звуковой дорожки watermarked- и reference- сигналов.

абсолютно необязательно. Эксперименты показали хорошие результаты с коэффициентом сжатие/растяжение порядка одного двух процентов, хотя это может увеличить риск возникновения различных искажений. Также отметим, что скорость кодирования порядка 8 бит в секунду реальна для аудио, и ограничена только теми недостатками, которые могут проявиться при заметной модификации временной оси.

Для многих приложений, таких как речь, можно использовать коэффициент сжатие/растяжение вплоть до 5 - 10 процентов, что соответственно увеличивает скорость кодирования.

Априорное рассмотрение

Этот метод имеет несколько значительных преимуществ над уже существующими. Во-первых, для большей части аудиозаписей изменения, вносимые в них этим методом, будут фактически не заметны, потому что слуховой аппарат человека не чувствителен к сильно низкой частотной модуляции. В то же время он исключительно устойчив к передаче и сжатию, потому что нынешняя цифровая аудио технология имеет точность времени порядка нескольких миллисекунд в час. Типичные аудио источники (речи или музыки) имеют достаточно естественные вариации (колебания), и такие, что искусственные "скоростные" изменения, внесенные кодированием водяного знака, будет трудно услышать. Исключением может быть весьма ритмичная музыка, созданная компьютерным синтезатором или другим механическим устройством. Далее, при тонком анализе промежутков такт за тактом можно обнаружить скоростные модификации, несмотря на то, что они могут быть незаметными для слушателя. Скоростные вариации, такие как те, что неотъемлемы в аналоговом записывающем и воспроизводящем оборудовании, не будут смешиваться с водяным знаком. То есть прямое скоростное изменение, например, из-за неточной скорости воспроизведения, не повлияет на водяной знак. Аналоговые записываемые неоднородности, обычно называемые "wow"(крик) и "flutter"(звуковая рябь), возникающие в шкале времени, значительно короче, чем скоростные изменения водяного знака, и поэтому они

в среднем не влияют на водяной знак при их воздействии на него. Возможно такое ограничение, что водяной знак может быть частично скрыт или размыт умышленным изменением шкалы времени аудио областей, или наложением скоростными изменениями другого водяного знака, но все это фактически не удаляет водяной знак, так как не известна точная инверсия использованных сжатий и растяжений (что требует опорного аудио для их определения). Заметим, что пока ни одна из любых других watermarking-схем не устойчива от приложения с многократных водяных знаков. ([2],[7])

Существующие методы создания водяных знаков для аудио

Типичные watermarking-методы для аудио можно классифицировать на методы в предметной области (области данных) и частотной области. Методы в предметной области фактически заключаются в модификации аудио данных, таких как "размывание" минимального количества важных битов в представлении кодово-импульсной модуляции или сокрытие данных с помощью сжатых областей. В методах частотной области модифицируют спектральное содержание сигнала, например, удаляя особенно редкую частотную компоненту или добавляя информацию, замаскированную под шум с малой амплитудой; также это могут быть такие методы: расширение спектра, фазовая модуляция, амплитудная модуляция и автокорреляционная модификация. Надо сказать, что могут быть и другие методы, которые не опубликованы и возможно являются торговым секретом. Метод модуляции оси времени достаточно новый, и он не укладывается ни в одной из этих категорий. Также важно, что этот метод имеет явные преимущества над обоими подходами: в отличие от методов в предметной области, он устойчив при передаче, кодировании и сжатии, и в отличие от методов в частотной области, не изменяет спектр и не добавляет шум. Действительно, модулирование масштаба времени совершенно "ортогонально" с типичными подходами: процесс создания водяного знака на основе "изменения" времени может проходить независимо, не пересекаясь с другими методами, и может быть объединен с любым из этих методов, чтобы увеличить информационную емкость и надежность водяного знака.

Методы в предметной области для создания водяных знаков включают:

- **Метод на основе сжатия областей.**

В этом методе водяной знак кодируется только в сжатое представление данных, и поэтому он в какой-то степени не устойчив. Когда данные не сжаты, водяной знак не доступен в отличие от нашего метода, где он сохраняется при аналоговом и цифровом копировании и при различных методах сжатия.

- **Размывание битов.**

Здесь информация кодируется модуляцией минимального числа важных битов при временно-областном или сжатом представлении. Хотя он потенциально и имеет большую скорость передачи данных, он не устойчив к сжатию, аналоговой передаче и воспроизведению, и вводит шум в сигнал.

- **Амплитудная модуляция.**

Максимальные уровни сигнала изменяются, чтобы попасть в заранее установленный амплитудный диапазон. Этот метод вводит модуляционное искажение, и не устойчив к амплитудному сжатию, которое широко используется в аналоговой и цифровой телефонии, радиовещании, усилении звука, уменьшении шума.

- **Echo hiding - Сокрытие эха.**

Отдельные копии сигнала смешиваются с оригинальным сигналом. Время задержки эхо достаточно короткое, и амплитуда копии тоже достаточно мала, чтобы все изменения были не слышны, хотя эхо может быть обнаружено через автокорреляцию. Этот метод вводит искажение, потому что имеется фазовое зануление на частотах, чей период кратен задержке эха. Также этот способ может быть не устойчивым при сжатии, так как незаметные эхо могут с большой вероятностью быть отброшены при кодировании.

Методы в частотной области для создания водяных знаков включают:

- **Фазовое кодирование.**

Этот способ полагается на то, что слуховой аппарат человека по отношению к фазе не чувствителен. Сигнал разбивается на окна как в спектрограмме, затем и модуль, и фаза вычисляются в каждом окне. Полностью искусственный фазовый сигнал, который кодирует водяной знак, вводится в первое окно. Фазовая информация для следующих окон(фреймов) многократно считается по фазовым различиям каждого фрейма с абсолютной фазой. Результирующая фаза объединяется с первоначальными модулями, чтобы окончательно закодировать водяной знак в сигнал. Этот метод вводит фазовый разброс в сигнал, и возможно он тоже не устойчив при сжатии.

- **Модификация полосы частот.**

Информация кодируется либо удалением или усилением определенных спектральных полос, либо удалением узкой спектральной частоты узкополосным режекторным фильтром, либо добавлением частотно-полосовых различий. Этот метод вводит спектральное искажение, он не устойчив к кодированию, и не работает также, если были изменены часто встречаемые в аудио частотные компоненты.

- **Spread spectrum - растяжение спектра.**

Сигнал, несущий информацию о водяном знаке, модулируется в широкополосный шум умножением с псевдослучайной последовательностью. Так как модуляционная функция известна (или может быть восстановлена), то сигнал водяного знака можно демодулировать. В этом методе добавляется шум к watermarked-сигналу, и низкая амплитуда сигнала с расширенным спектром означает, что шум, вероятно, может быть отброшен при кодировании. К тому же частота дискретизации широко используется как несущая частота при модуляции для синхронизации с получателем. В этом случае повторная дискретизация или аналоговая передача, вероятно, "уничтожат" синхронизацию вместе с водяным знаком.

В итоге можно сказать, что схемы создания водяных знаков с помощью сжатия областей мало используются, потому что они не работают, тогда как модуляция оси времени работает, в несжатом аудио сигнале. Много других схем, особенно "размывание" и методы в частотных областях, не устойчивы к сжатию аудио. Это особенно проблематично в отношении хорошего водяного знака, когда частотные модификации должны быть не слышимы, однако точно известно, что информация теряется или изменяется, когда используются такие схемы сжатия как MP3. В данном случае наш метод устойчив и при таком кодировании, и при аналоговом копировании; и в отличие от других методов, он не вводит шум, эхо, спектральное или коротко временное фазовое искажение в watermarked-сигнал, к тому же не удаляет или иначе не изменяет частотные компоненты или амплитуды сигнала.

Методы создания водяных знаков для изображений

Существуют большое число работ в кодировании водяных знаков в изображение. Большинство подходов во многом похожи на те, что были описаны выше, например методы с растяжением спектра могут быть применены для изображений так же, как и для аудио. Один важный типичный способ создания водяного знака для текста — это модулирование пробелов между словами и предложениями. В этом методе требуется определить границы слов, и он не применим к общим не текстовым изображениям. Glyph технология Xerox PARC кодирует информацию в цифровую документальную копию, используя крошечные отметки, которые могут быть размыты при создании "серых теней". Метод, называемый текстурно-блоковым кодированием, кодирует информацию копированием областей произвольной текстуры. Эти области можно найти автокорреляцией. Как с аудио, эти методы все "ортогональны" к методу модуляции оси времени (для изображений, лучше для него название — пространственная модуляция), и они могут быть использованы во взаимодействии.

Создание водяного знака "модуляцией времени"

Способы масштабирования основного тона аудио сигнала хорошо известны и находятся в широком использовании. Эти методы можно использовать в равной степени хорошо для изменения длины аудио записи без нежелательной модификации основного тона, вводимой простым изменением скорости. Общий метод масштабирования времени основан на кратковременном Фурье преобразовании, но другие методы, такие как Phase Vocoder, Time Domain Harmonic Scaling и Pitch-Synchronous Overlap Add (PSOLA), также широко используются. Любой метод масштабирования времени можно использовать для создания водяного знака; предпочтительные методы могут сжимать или растягивать с коэффициентами очень близкими к единице с мало слышимым искажением.

Рассмотрим сигнал $x(t)$, выраженный как объединение K не пересекающихся блоков $x_{1...K}$ с конкатенацией, обозначенной \mathbf{C}

$$x(t) \equiv \mathbf{C}_{k=1}^K x_k$$

Watermarked-сигнал, $x_w(t)$, генерируется при выполнении сжатия или растяжения каждого блока с коэффициентом E_k , и объединяется в модифицированные блоки

$$x_w(t) = \mathbf{C}_{k=1}^K TCM(x_k, E_k).$$

Где $TCM(x_k, E_k)$ обозначает временную модификацию блока x_k в E_k раз. На практике требуется внимательность, чтобы избежать введения слышимых сосредоточенных неоднородностей на границах блоков. Это можно выполнить, используя определенный алгоритм, который оставляет данные на или около границ блоков неизменными, или перекрывает слегка сегменты и усредняет данные внутри области перекрытия во время создания watermarked-сигнала.

Последовательность величин E_k , растяжения/сжатия, кодирует водяной знак, и его можно обнаружить, сравнивая watermarked-версию с опорным аудио. Это делается нахождением функции, трансформирующей время и противопоставляющей опорный файл с watermarked-файлом. Вычитание линейной компоненты выровненного пути дает "скоростную" карту, по которой можно определить информацию о водяном знаке. После удаления линейной компоненты, траектория имеет положительный наклон в сжатых областях и отрицательный в растянутых областях, и наклон ноль в неизменных областях. Это проиллюстрировано на рисунке 3. Наклон обнаруженной скоростной карты соответствует величинам E_k , использованным во время кодирования. Этот факт будет использован ниже для обнаружения водяного знака.

В потоковой реализации сопоставляют опорный и watermarked- сигналы, используя коротко временные свойства модуля. Если анализируемые фреймы длиной по 128 отсчетов, то для аудио с частотой дискретизации 22.05 kHz, это приводит к 5.8 ms фрейма и фреймовой скорости 172 за секунду. Каждый анализируемый фрейм "окнируется" с помощью оконной функции Хэмминга и преобразуется в частотную область быстрым Фурье преобразованием (FFT). Логарифм модуля результата используется как оценка спектральной функции фрейма. Результирующий вектор спектральных компонент характеризует спектральное содержимое окна. Последовательность спектральных векторов изображает частотное содержание сигнала со времени. Некоторыми частотными компонентами можно по выбору пренебречь, если они не такие полезные для подобных вычислений.

Чтобы найти наилучший деформационный путь, вычисляются спектрограммы и для опорного, и тестового сигналов. Спектрограммы выравниваются деформацией опорного, чтобы сопоставить с тестовым сигналом. Это делается с использованием динамического программирования (DP). Показано, что оптимальный выравнивающий путь вычисляется за квадратичное время, то есть время пропорционально произведению длин сравниваемых последовательностей. DP алгоритм особенно хорош при специальных условиях этого приложения. Например, DP изящно работает в случае, когда тестовый и опорный файлы не начинаются и не заканчиваются в точно одинаковое время. Например, если watermarked-сигнал был извлечен из непрерывной радиотрансляции, то не нужно определять начальные

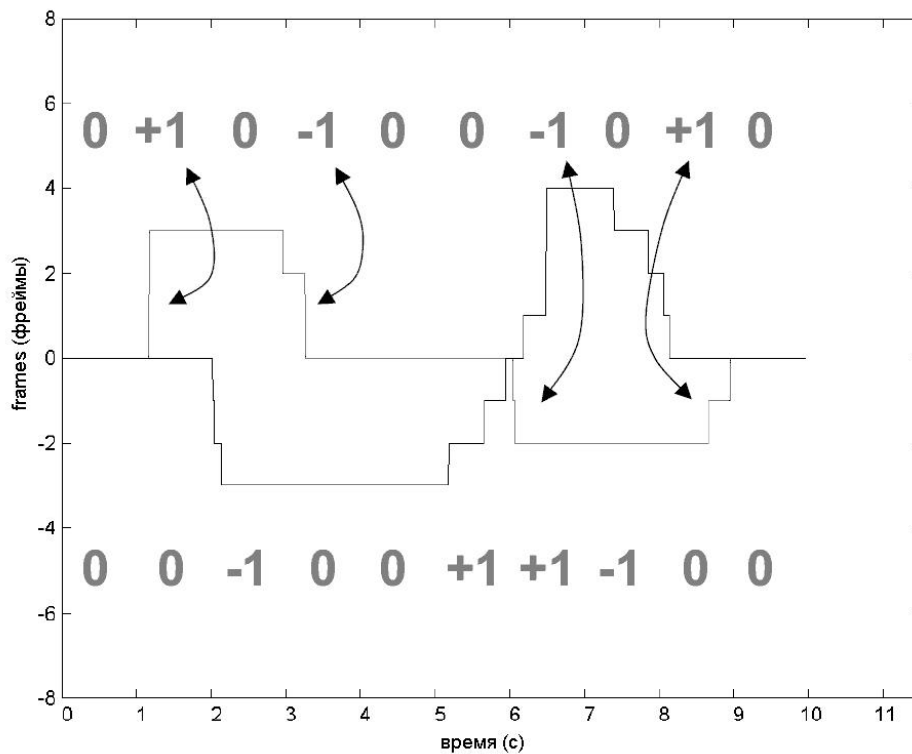


Рисунок 3. Два разных 'watermarks', восстановленных динамическим программированием (DP). Сжатые области обозначаются +1, а растянутые -1.

и конечные точки точно. Другое свойство это то, что спектру фрейма необязательно совпадать точно с фреймом опорного сигнала, так как они будут более похожими, чем их соседи, таким образом, система устойчива к разумным спектральным искажениям. Также в этом приложении ожидаемое смещение очень маленькое; то есть лучший путь незначительно отклоняется от диагонали. В этом случае DP алгоритм может действительно сделать линейным время при "отбрасывании углов", то есть, считая только пути очень близкие к диагонали. Подобным образом, полное время модификации (вызванное, например, частотой отсчетов преобразования или ошибочной, некорректной аналоговой скоростью воспроизведения) изящно регулируется при DP алгоритме. В этом случае водяной знак восстанавливается вычитанием диагонали прямоугольника, сформированного при комбинационном произведении двух сигналов, а не просто как при возведении в квадрат.

Суммарная скорость данных водяного знака — это компромисс между способностью обнаружить его и ухудшением сигнала. Назовем минимум длины интервала сжатия/растяжения блоком. Для простоты, предположим, что все блоки имеют одинаковую длину (хотя не сжатые блоки могут быть и меньше, чем измененные блоки). Каждый блок может быть сжат или растянут в $1 \pm \epsilon$ раз (коэффициент). Если ϵ достаточно мало, тогда сжатие/растяжение можно продискретизировать интегральным фактором, то есть $1 \pm n\epsilon$, посредством n малого целого, возможно отрицательного. Блоки могут оставаться несжатыми, то есть $n = 0$. Чтобы ослабить слышимые помехи, целесообразно (хотя жестко не необходимо), чтобы величина n была ограничена, а именно, меньше чем маленькая величина N . По той же причине, изменение в n должно быть маленьким между расположенными рядом блоками. Чтобы сохранить временную длину файла, предпочтительно, хотя снова необязательно является необходимым, чтобы сумма n была равна нулю по всем блокам, так что количество сжатий точно равно количеству растяжений. Таким образом, каждый блок k имеет прикрепленный множитель, n_k , такой, что $N \leq n_k \leq N$ и соответствующее значение сжатия/растяжения $E = 1 + n_k\epsilon$. Конкретному водяному знаку соответствует последовательность E_0, E_1, \dots, E_K . Эту последовательность можно впоследствии восстано-

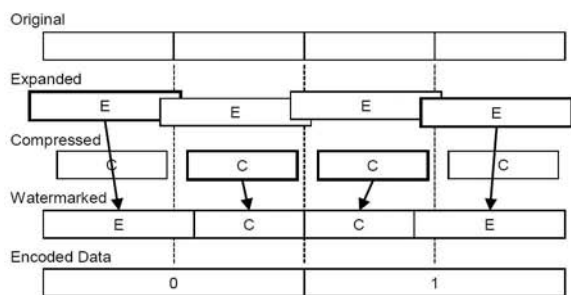


Рисунок 4а. Конструкция 2-х битного watermarked-сигнала из сжатых и растянутых частей.

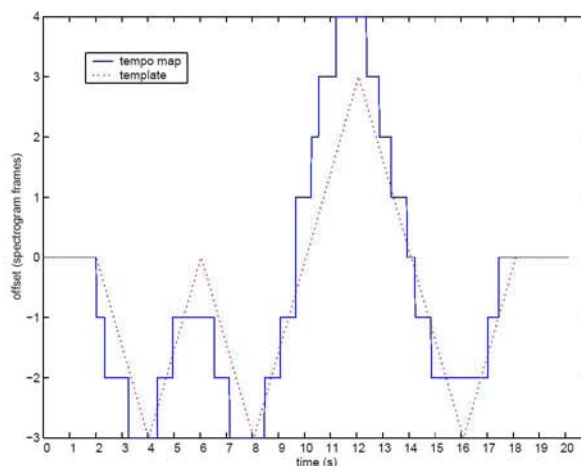


Рисунок 4б. Восстановленная 'tempo map' и ожидаемый шаблон для двоичного '0010' 'watermark'.

вить квантованием производной скоростной карты. Эксперименты показали, что разумные величины блоковой длины 0.5 секунд, $\epsilon = 0.01$ и $N = 2$. Так каждую секунду аудио можно закодировать примерно $2 \log_2 (2N + 1)$ битами, или незначительно больше, чем 8 бит. Пока скорость данных не огромная, заметим, что типичная популярная песня 180 секунд длины. Скорости данных 8 бит в секунду достаточно, чтобы закодировать 180 байтов, или более чем достаточно для названия песни, артиста, издателя, и ID. Когда 180 байтов используются как водяной знак, это дает более чем 10^{400} индивидуальных идентификаторов, которых бы было более чем достаточно для любой мыслимой комбинации идентификаторов источника, устройства, и времени создания.

Рисунок 3 показывает результаты демонстрации создания примитивного водяного знака. В две копии 10 секунд отрывка песни Битлз были закодированы водяные знаки. Коэффициенты сжатия и растяжения были по два процента, над односекундными областями, так общее смещение 20 ms, или 3.44 фрейма. Первая копия была сжата в 1 и 8 секундах и растягнута в 3 и 6 секундах. Другая копия была растягнута в 2 и 7 секундах и сжата в 5 и 6 секундах. Скоростные карты из алгоритма динамического программирования показаны на рисунке 4б, которые вычерчивают отклонение от линейного времени в спектрограммных фреймах. Области сжатия и растяжения легко обнаружимы, как плата начала времени смещения. Этот метод был испытан для других аудио областей, включающих и звуковые разговорные дорожки, и оркестровую музыку, с одинаково хорошими результатами. Свободные тесты по прослушиванию из 13 добровольных слушателей подтвердили "не слышимость" модуляции оси времени, хотя один "golden-eared" слушатель был способен обнаружить искажения по временному сжатию при прямом А/В сравнении с опорным аудио. Наилучшие методы сжатия времени возможно имеют меньшее число искажений. Дополнительный эксперимент продемонстрировал, что водяной знак легко выживает при 64 кВ MP3 кодировании и декодировании.

Другой эксперимент проверил восстанавливаемость водяных знаков. В этом случае, опорным сигналом были первые 20 секунд песни "Magical Mystery Tour" Битлз, конвертированные в монофоническое представление на частоте дискретизации 20,050 Hz. (Метод будет, конечно, работать с любой частотой дискретизации и может быть применен к стерео или мультисканальному звуку).

Простое кодирование было использовано, чтобы закодировать водяной знак длиной 4 бита в 16 различных копиях аудио. В этой схеме использовали два двухсекундных блока, каждый кодирует один бит информации. Сжатие, за которым следует растяжение, представляет бинарную единицу, тогда как в обратном порядке это ноль. Посмотрите на рисунок 4а. Очевидно, что существует достаточное число кодирующих схем, особенно тех, что ис-

пользуют область без модуляции оси времени, чтобы кодировать дополнительное состояние.

Если дана эта кодирующая схема, то сгенерировать watermarked-сигнал налету просто, сцепляя сжатые и растянутые области сигнала. В этом случае, сжатые и растянутые версии сигнала были созданы с коэффициентом 2.5%; таким образом, растянутая версия имела длительность 20.5 секунд. Каждая версия была равномерно разделена на 10 равных блоков, каждая по две секунды длиной. Watermarked-сигнал был создан простым методом сцепления сжатых и растянутых блоков. Блоки в начале и в конце сигнала были не сжаты, так только средние 16 секунд были изменены. (В определенных случаях, возникают, как следствие, слышимые искажения из-за плохо сочетающихся блоковых краев; в практической системе они могут быть устранены перекрестным размытием блоков за короткий период.)

Если известна последовательность сжатия и растяжения, то это сразу ведет к установлению, какой скоростная карта должна быть в каждом случае. Переход от области сжатия к области растяжения соответствует на скоростной карте росту скорости, а затем замедлению, указывая на бинарную единицу. И наоборот, переход от растяжения к сжатию изображается формально сигналом ниже нуля в форме "V". Скоростная карта есть отклонение от скорости сжатия/растяжения. Так как существуют константы для образов блоков, результирующую скоростную карту можно предсказать пилообразным сигналом. Соответственно, "шаблоны" строятся на основе пилообразных выбросов, соответствующих ожидаемым скоростным изменениям. Рисунок 4а показывает шаблон для watermarked- копии с бинарными "0010". Это согласуется с последовательностью сжатие/растяжение СЕСЕЕССЕ, где "С" и "Е" представляют сжатые или растянутые области двухсекундного блока. Скоростная карта была извлечена для каждой watermarked-версии, и сравнена с 16 ожидаемыми шаблонами. Если известны и скоростная карта \vec{m} , и шаблон \vec{t} , то полезной метрикой будет косинус угла между ними

$$D_c(\vec{m}, \vec{t}) \equiv \frac{\vec{m} \cdot \vec{t}}{\|\vec{m}\| \cdot \|\vec{t}\|}.$$

Он обладает таким свойством, которое дает хорошую оценку подобия, не обращая внимания на абсолютные векторные величины. Эксперимент показал, что ожидаемые водяные знаки показывают намного больший косинус между ожидаемым шаблоном, чем любым другим шаблоном. Из 256 сравнений шаблонов скоростных карт, максимумом косинусного расстояния для неправильного шаблона является 0.618, тем временем как минимум оценки для правильного шаблона 0.9094. Все правильные шаблоны из 256 имели оценки более 0.9, тогда как все неправильные шаблоны имели оценки меньше, чем 0.62. Разница в оценках пропорциональна расстоянию Хэмминга между водяным знаком и шаблоном. Чтобы увеличить расстояние оценки, можно использовать подмножество кодов с большими расстояниями Хэмминга. Например, если использовать 8 четырех битовых кодов с четной суммой (гарантируя расстояние Хэмминга два или более), то максимум оценки неправильного шаблона уменьшается к 0.238. Отметим, что необязательно использовать шаблоны; водяной знак может быть извлечен сразу же порогово-классификационными методами. Например, исследование рисунка 4 показывает, что бинарные цифры могут быть извлечены, если пределы установить в ± 1 фреймы.

Использование нелинейной масштабной модификации для создания водяных знаков в изображениях и отсканированном тексте

Помимо аудио, этот метод можно использовать для создания водяных знаков еще и в изображениях. Как и в случае с аудио, этот метод имеет преимущество, он устойчив при сжатии с потерями и аналогом копировании. К тому же, кодирование водяных знаков можно легко выполнить как по "цифровому", так и даже напрямую через механизм принтера или ксерокса. Это может быть особенно ценно для высоко секретных приложений: ксерокс или принтер могут кодировать время, дату, и устройство или user ID невидимо в каждой

сделанной копии или распечатке. Если незаконное копирование обнаружено, водяной знак может помочь определить источник.

В этом методе участки изображения сжимаются или растягиваются в незаметном количестве. Хорошо известные цифровые методы на основе повторной выборки изображения с изменением его разрешения могут вытягивать или сжимать области изображения на малое количество. В качестве альтернативы можно использовать механические или оптические методы, такие как изменение скорости барабана или валика сканера, изменение печатающей скорости головки принтера, или изменение скорости объективной цилиндрической линзы в ксероксе по отношению к барабану. По крайней мере доступны две оси деформирования (вертикальная и горизонтальная); возможно дифференцированно деформировать "полосы" в изображении, хотя это может привести к заметным искажениям, так как прямые линии не параллельные осям деформации будут больше совсем не прямыми. Для небольшого класса изображений это может привести к видимым искажениям, особенно на изображениях, которые имеют постоянные линии или решетки, идущие по диагоналям. Как в аудио случае, изображение можно проанализировать для нахождения областей или формы водяного знака, которые внесут наименьшее число заметных изменений. Например, применяя Фурье анализ изображения, можно найти угловое направление с наименьшей энергией. Если использовать это направление как ось деформации, то заметные искажения будут минимизированы. Например, если дано изображение параллельных линий, Фурье анализ может легко найти направление линий. Деформация изображения параллельно этому направлению даст результат с наименее заметными изменениями. В общем, если задана малая степень деформации, кодирование водяного знака в большом количестве изображении будет не ощутимо для них.

Чтобы восстановить водяной знак, опорное изображение сравнивается с watermarked-изображением. В отличие от аудио, пиксели можно сравнивать напрямую. Столбцы пикселей перпендикулярных к оси деформации можно сравнить по Евклидовому или другому расстоянию, как например, спектральные вектора сравниваются в аудио случае.

.....

Особенное преимущество этого метода состоит в том, что он вычислительно приемлемый для кодирования водяного знака на клиентской стороне, являясь частью streaming media player. Клиент затем разжимает и расшифровывает streaming media и применяет водяной знак до буферизации, проигрывания и локального сохранения сигнала, таким образом, уменьшая вычислительные нагрузки на сервер. Действительно, водяной знак мог бы кодировать информацию относительно пользователя, тогда в случае утечки информации, пользователя можно бы было идентифицировать по этим данным.

Недостаток этого и многих типичных watermarking-методов - это необходимость опорного сигнала, чтобы раскодировать водяной знак. Несмотря на то, что он отлично подходит для многих приложений, существуют приложения, где он должен извлекать водяной знак без опорного файла. Создание водяного знака на основе модуляции времени может быть использовано при отсутствии опорного сигнала, если реальная скорость может быть предварительно оцененной или предсказанной. Существует много методов для анализа скорости (темпа) или скорости разговора в аудио. Из всего это возникает вопрос, как сконструировать "предсказатель" скорости, которые бы оценивал сигнал на некоторый короткий период времени в будущем...

Литература:

1. J. Zhao, E. Koch and C. Luo, "Digital Watermarking In Business Today and Tomorrow", in Communications of ACM. Vol. 41, No. 7, July 1998.
<http://syscop.igd.fhg.de/pubs.html>
<http://syscop.igd.fhg.de/Publications/Zhao98b.pdf>
2. Michael Arnold, Zongwei Huang, "Blind detection of multiple Audio Watermarks", Department for Security Technology for Graphics and Communication Systems Fraunhofer-Institute for Computer Graphics
<http://syscop.igd.fhg.de/Publications/Arnold01a.pdf>
3. S. Sprenger, "Time and Pitch Scaling of Audio Signals",
<http://www.dspdimension.com/html/timepitch.html>
4. D. Dorran, R. Lawlor and E. Coyle, "High quality time-scale modification of speech using a peak alignment overlap-add algorithm (Paola)", Dublin Institute of Technology, National University of Ireland, Maynooth, 2003
<http://nlp.iit.net/~speech/rohit/papers/01-00700.pdf>
5. I. Cox, J. Kilian, T. Leighton and T. Shamoan, "Secure Spread Spectrum Watermarking for Images, Audio and Video", in Proc. of 1996 Int'l. Conf. on Image Processing (ICIP'96), vol. III, pp. 243-246, (1996).
<http://clip.informatik.uni-leipzig.de/~toelke/Watermark/0676.pdf>
6. J. Foote and S. Uchihashi, S., "The Beat Spectrum: A New Approach to Rhythm Analysis", in Proc. IEEE International Conference on Multimedia and Expo (ICME) 2001
<http://www.fxpal.com/people/foote/papers/ICME2001.htm>
7. Mohamed F. Mansour and Ahmed H. Tewfik, "Audio Watermarking by Time-Scale Modification", Department of Electrical and Computer Engineering, University of Minnesota
<http://www.ece.umn.edu/users/mmansour/AudioWM.PDF>