

УДК 612.85

## ОСОБЕННОСТИ ФОРМИРОВАНИЯ ВИРТУАЛЬНЫХ АКУСТИЧЕСКИХ ОБЪЕКТОВ

© 2000 г. А. В. Аграновский, Г. Е. Евреинов

ГП КБ "СПЕЦВУЗАВТОМАТИКА"

344007 Ростов-на-Дону, Газетный, 51

E-mail: asni@ns.rnd.runnet.ru

Поступила в редакцию 24.02.97 г.

Как происходит восприятие озвученного контурного изображения? Какие дополнительные признаки необходимо установить, чтобы стабилизировать восприятие границ виртуального звукового объекта и сформировать слуховое ощущение, изоморфное зрительному? В работе обсуждается возможность коррекции восприятия субъективных образов путей управления параметрами (девиация частоты и длительность критических точек) виртуальных акустических объектов, формируемых иллюзорным движением звуковых источников.

В отличие от произвольно фиксируемого зрительного внимания, при формировании акустических объектов путем иллюзорного движения виртуальных звуковых источников, необходимо эффективно управлять теми параметрами, с помощью которых возможно постоянно удерживать "фокус" внимания слухового анализатора на разворачивающейся во времени звуковой последовательности. Процесс управления вниманием очень сложен, а в виртуальных средах до сих пор недостаточно изучен.

Как происходит восприятие озвученного контурного изображения? Какие дополнительные признаки необходимо установить, чтобы стабилизировать восприятие границ виртуального звукового объекта и сформировать слуховое ощущение, изоморфное зрительному?

Для формирования виртуальных звуковых объектов мы используем векторный аудиодисплей [1, 2, 3]. Устройство состоит из одноместного звукового модуля, управляемого с помощью специального интерфейса, который обслуживает четыре электродинамических акустических излучателя с обрамлением, ограничивающим сектор озвучивания до 90 угл. град., установленных в углах квадрата со стороной 0.46 м и на расстоянии около 0.6 м от слуховых каналов испытуемого, голова которого свободна (инструкций на ограничение движений не устанавливалось), глаза закрыты. Расстояние 0.46 м принято за основание и высоту акустической плоскости. Программный интерфейс позволяет создавать графические объекты на экране монитора, которые затем озвучиваются движением одного или нескольких звуковых курсоров (виртуальных звуковых источников), формируемых путем управления пространственным положением интерференционно-

го максимума звуковых сигналов четырех акустических излучателей по закону:

$$A_1 + A_2 + A_3 + A_4 = B = \text{const} \quad (1)$$

или

$$AX(1 - Y) + A(1 - X)(1 - Y) + AX Y + A(1 - X)Y = B \quad (2)$$

(где  $A$  – амплитуда сигнала,  $X, Y$  – текущие координаты, в интервале  $[0, 1]$ , причем левая нижняя точка плоскости изображения есть начало координат,  $B$  – субъективный уровень громкости звукового курсора). Таким образом, можно сравнивать субъективный звуковой образ, описываемый испытуемым с реальным изображением на мониторе экспериментатора. Кроме того, виртуальную акустическую плоскость имитировали с помощью листа бумаги, на котором испытуемый воспроизводил фломастером субъективные впечатления сразу после прослушивания предъявленных звуковых объектов, глаза при этом оставались закрытыми.

Общий вид части интерфейса, связанной с разработкой графических объектов, показан на рис. 1А. Он состоит из программных модулей, реализованных в среде Windows'95 на языке Visual Basic 4.0. Модуль формирования графических объектов имеет рабочее поле с фиксированным числом точек ( $M$ )  $128 \times 128$ . В некотором смысле оно аналогично матрице виртуальных звуковых источников [4, 5, 6], но имеет в восемь раз более высокое пространственное разрешение и обеспечивает необходимую непрерывность озвучивания без применения специальных методов интерполяции звукового сигнала между точками. Интерпо-

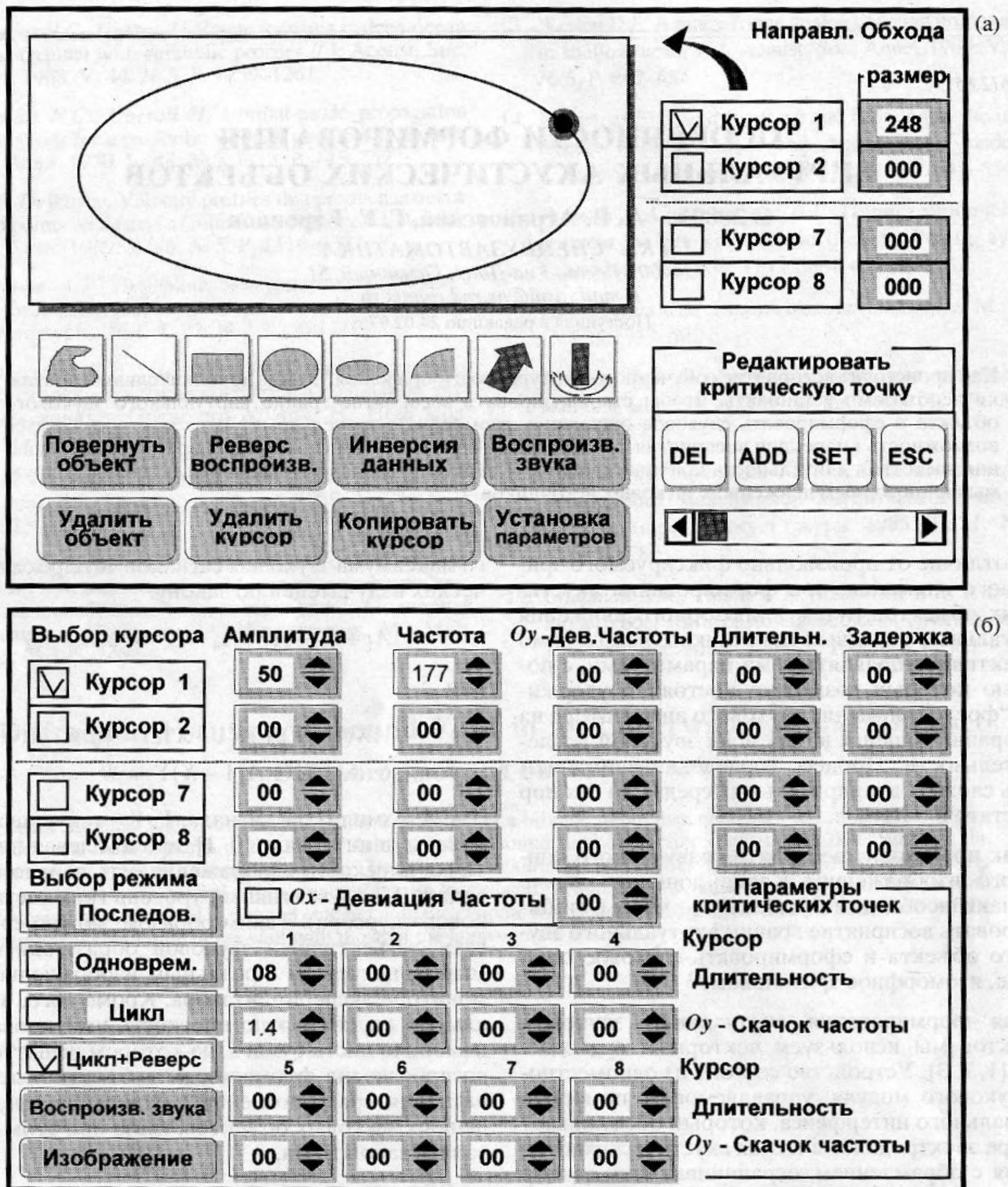


Рис. 1. Проектирование графических объектов (а) и управление звуковыми параметрами (б).

ляция используется только при построении изображения.

Панель инструментов предназначена для формирования стандартных графических объектов: прямая линия (отрезок, ломаная), дуга, окружность, эллипс, прямоугольник, произвольная кривая линия, и возможности их копирования или произвольного перемещения в пределах рабоче-

го поля. Панель управления позволяет редактировать графические объекты относительно их звуковых атрибутов. Имеется дополнительная панель поточечного редактирования фрагментов изображения. Эта функция связана с маркировкой критических точек (крайних или промежуточных), необходимых для управления процессом восприятия, и установкой произвольного числа незавучающих точек.

Общий вид звуковой части программного интерфейса показан на рис. 1б. Изменение частоты звукового сигнала регулируется в пределах 100–5000 Гц, но в реальных экспериментах не превышало 800 Гц. Встроенные генераторы и аттенюаторы звукового модуля, управляемые с помощью восьмиразрядных цифроаналоговых преобразователей, обеспечивали конечное целое число частот и уровней амплитуды. Относительно широкая полоса звукового сигнала обеспечивалась его формой, близкой к прямоугольной. Регулировка девиации частоты  $D_x$  и  $D_y$  осуществлялась в зависимости от значения координат точки изображения  $(X, Y)$ , для каждого курсора независимо по оси  $Oy$  и для всех курсоров одновременно по оси  $Ox$  при установленном начальном значении частоты  $F_0$  в положении  $X = 0, Y = 128$ . То есть значение частоты понижалось, если точка изображения перемещалась сверху вниз и повышалось при движении этой точки слева направо так, что самая низкая частота оказывалась в левой крайней позиции нижней части виртуальной плоскости, а самой высокой по частоте являлась правая верхняя точка этой плоскости.

Дополнительно устанавливаются параметры длительности звучания и скачка частоты в критических точках для каждого фрагмента звукового объекта (крайних или промежуточных).

Временные параметры устанавливаются отдельно по каждому курсору с дискретностью 1 мс. Следует различать параметр, обозначенный как “длительность”, соответствующий скорости изменения значений текущих координат по заданному курсору (что эквивалентно длительности звучания текущей точки массива), и параметр, обозначенный как “задержка”, соответствующий разным временным параметрам в зависимости от режима воспроизведения звукового объекта. При последовательном воспроизведении звуковых курсоров “задержка” осуществляет отсроченное воспроизведение первой точки следующего фрагмента относительно последней точки предыдущего. При одновременном воспроизведении звуковых курсоров “задержка” отодвигает момент воспроизведения первой точки предъявляемого фрагмента от момента нажатия на кнопку “звук”, причем включение воспроизведения звукового объекта может быть оставлено на постоянную величину. Длительность звучания критических точек определяется как число, кратное длительности звучания текущей точки плюс 4 мс.

Озвучивание фрагментов изображения, сформированных соответствующими звуковыми курсорами, может осуществляться в последовательном или одновременном режиме, с синхронизацией по моменту запуска или асинхронно по каждому курсору, однократно, с возможностью реверсирования или циклически.

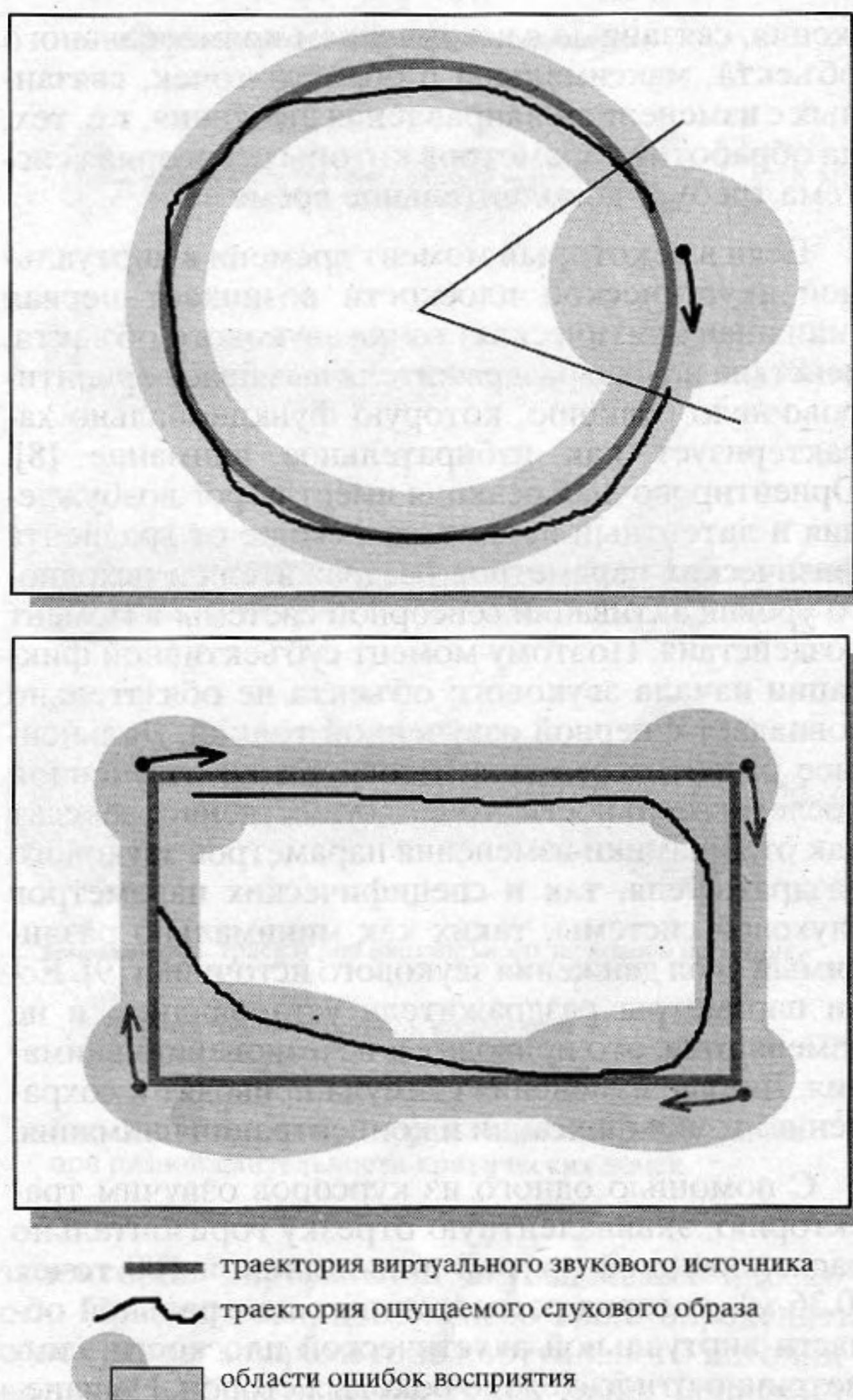


Рис. 2. Восприятие виртуальных звуковых объектов в отсутствие маркеров критических точек.

Программный интерфейс обеспечивает стандартный сервисный набор подпрограмм работы с созданными файлами.

Исследуем восприятие озвученного фрагмента изображения. Для большей наглядности выберем замкнутые фигуры, например: окружность и прямоугольник. На рис. 2 показаны исходные графические объекты и их субъективные слуховые образы по результатам графических отчетов испытуемых при однократном воспроизведении и в отсутствие специальных признаков выделения начальной точки объекта. Для окружности, в области начальной точки можно указать сектор, в котором неопределенность локализации траектории приводит к существенным ошибкам восприятия предъявленного звукового объекта. Величина этого сектора зависит от скорости движения виртуального источника, радиуса окружности и спектральных параметров звукового сигнала. Иско-

жения, связанные с восприятием прямоугольного объекта, максимальны в области точек, связанных с изменением направления движения, т.е. тех, на обработку параметров которых сенсорная система требует дополнительное время.

Если в некоторый момент времени в виртуальной акустической плоскости возникает первая (или иная критическая) точка звукового объекта, действие нового раздражителя вызывает ориентировочную реакцию, которую функционально характеризует как избирательное внимание [8]. Ориентировочная реакция имеет порог возбуждения и латентный период, зависящие от градиента физических параметров раздражителя и исходного уровня активации сенсорной системы в момент воздействия. Поэтому момент субъективной фиксации начала звукового объекта не обязательно совпадает с первой озвученной точкой. Дальнейшее развитие реакции и восприятия озвученной последовательности точек существенно зависит как от динамики изменения параметров звукового раздражителя, так и специфических параметров слуховой системы, таких как минимально различимый угол движения звукового источника [9]. Если параметры раздражителя установились, и не изменяются, это приводит к исчезновению внимания. Любые изменения стимула приводят к сохранению точки фиксации и концентрации внимания.

С помощью одного из курсоров озвучим траекторию, эквивалентную отрезку горизонтально расположенной прямой линии длиной 100 точек (0.36 м), которая расположена в центральной области виртуальной акустической плоскости, симметрично относительно боковых сторон. На панели задания звуковых параметров значение частоты основного тона установим 305 Гц. Девиацию частоты по осям  $Oy$  и  $Ox$  и параметры дополнительных маркеров критических точек установим в неактивное состояние. Тогда, субъективное восприятие длины заданного отрезка при однократном воспроизведении будет зависеть от длительности звучания 100 точек (последовательно возбуждаемых виртуальных источников).

Установим амплитуду звукового курсора на уровне 50 дБ. При длительности звучания каждой точки 4 мс, мощности виртуального источника звука (частота 305 Гц) в первой точке может оказаться недостаточно, чтобы надежно локализовать начало отрезка. Для этого может понадобиться воспроизведение нескольких точек. Такой же эффект может наблюдаться во время воспроизведения всего отрезка, в том числе и в его конечной точке. Поэтому оценка длины предъявляемого виртуального звукового отрезка в такой постановке эксперимента малоэффективна. Дисперсия длин траекторий в отчете испытуемых может быть соизмерима с длиной самого отрезка при любой методике регистрации, а число точек

внутри отрезка, так же, как и частота основного тона малосущественны.

Заметим, что дискретизация виртуальной траектории (скачкообразные изменения амплитуды четырех излучателей каждые 4 мс) может вызывать ощущение "вибрато". Но этот эффект наблюдается лишь в случае использования синусоидальных сигналов и числе точек  $M < 64$  (3), а также соизмеримых значениях амплитуды тона и модуляции. Скачкообразные изменения уровня прямоугольного сигнала (с частотой 250 Гц), "сопровождающие" перемещение звукового курсора (305 Гц) имеют величину не более 0.5 дБ (при  $M = 128$ ). Интерференция таких сигналов практического значения не имеет, так как разностная частота лежит на границе как слухового восприятия, так и частотного диапазона используемых акустических систем. Поэтому специальные меры по обеспечению непрерывной формы звукового сигнала при  $M \geq 128$  не требуются.

Результаты экспериментов [6] по распознаванию алфавитно-цифровых символов также указывают на необходимость выделять начальную точку каждого звукового объекта. В ходе такого эксперимента в этом случае получалось 60–90% верных ответов. Еще большее значение имеет выделение критических точек при взаимодействии нескольких звуковых курсоров в одном звуковом объекте (рис. 6). Но увеличение длительности звучания точки от 5 мс до 16 мс изменяет субъективное ощущение длины отрезка (100 точек) в два раза. Увеличение длительности приводит к диффузному восприятию крайних точек, уменьшение – к вырождению впечатления протяженного объекта в диффузную область.

Сначала, для маркировки критических точек, мы использовали длительность звучания: повторное озвучивание точки (без изменения пространственной позиции) заданное число раз. Но это частое решение оказалось мало эффективным при увеличении сложности и скорости воспроизведения звуковых объектов. Длительность звучания критических точек, достаточная для стабилизации внимания, составляла от 40 до 200 мс. При длине отрезка 100 точек, на крайние точки приходилось до 50% времени звучания.

Так как в будущих экспериментах планировалось исследование возможности формирования квазистатических виртуальных звуковых объектов (курсор, масштабные линейки, границы окон и другие атрибуты графических интерфейсов), в звуковой части программного интерфейса был введен режим циклического воспроизведения звукового объекта с реверсированием направления воспроизведения. При определении оптимальных значений некоторых частотных и временных параметров, стабилизирующих восприятие звукового объекта, подбор значений осуществлялся в

циклическом режиме с реверсом, но результат проверялся при однократном воспроизведении. Например, при установленной частоте и длительности воспроизведения точки 4 мс, в отсутствие специального выделения критических точек, циклическое воспроизведение отрезка прямой линии вызывало ощущение звуковых биений, но эффект горизонтального движения ощущали не все испытуемые (8 человек). При удлинении времени звучания крайних точек до 40 мс отчетливо различались две диффузные области, между которыми ощущалось движение звука. При длительности звучания крайних точек 60 мс впечатление движения виртуального источника становилось более отчетливым, но возрастала дисперсия восприятия крайних позиций, которая уменьшалась при увеличении времени звучания до 8 мс в каждой точке.

Поэтому мы решили исследовать возможность маркировки критических точек путем одновременного изменения длительности их звучания и частоты основного тона – скачок частоты в критической точке.

При тех же условиях (100 точек, 305 Гц) для выделения крайних позиций горизонтального отрезка прямой линии минимальная длительность озвучивания крайних точек была установлена равной 8 мс, а частота звука на это время изменялась скачком от 305 Гц до 725 Гц. Эта величина во много раз превышала значение, необходимое для восприятия и эффективной пространственной локализации критической точки при установленной длительности ее воспроизведения, и вызывала устойчивое и примерно одинаковое ощущение у всех испытуемых.

Оказалось, что при увеличении длительности звучания крайних точек отрезка (от 8 до 20 мс и более) и одновременном изменении частоты, наблюдались существенные отклонения формы отрезка от линейной. На рис. 3 схематически показаны субъективные образцы предъявленной звуковой модели отрезка прямой линии при различных длительностях звучания крайних точек и значений девиации частоты  $D_y = 1.3$ .

Так как 8 мс длительности скачка частоты было достаточно для удовлетворительной локализации точки (вероятность обнаружения  $P > 90\%$ ), мы решили исследовать зависимость необходимого скачка частоты в крайней точке от частоты основного тона движущегося горизонтально курсора (при отсутствии девиации частоты основного тона по осям:  $D_y = 1.0$ ,  $D_x = 1.0$ ). Эта задача решалась путем измерения субъективного порога восприятия кратковременного изменения частоты движущегося виртуального источника прямоугольного сигнала при заданной частоте, скорости перемещения источника, а также длительности и локализации изменения в заданной точке

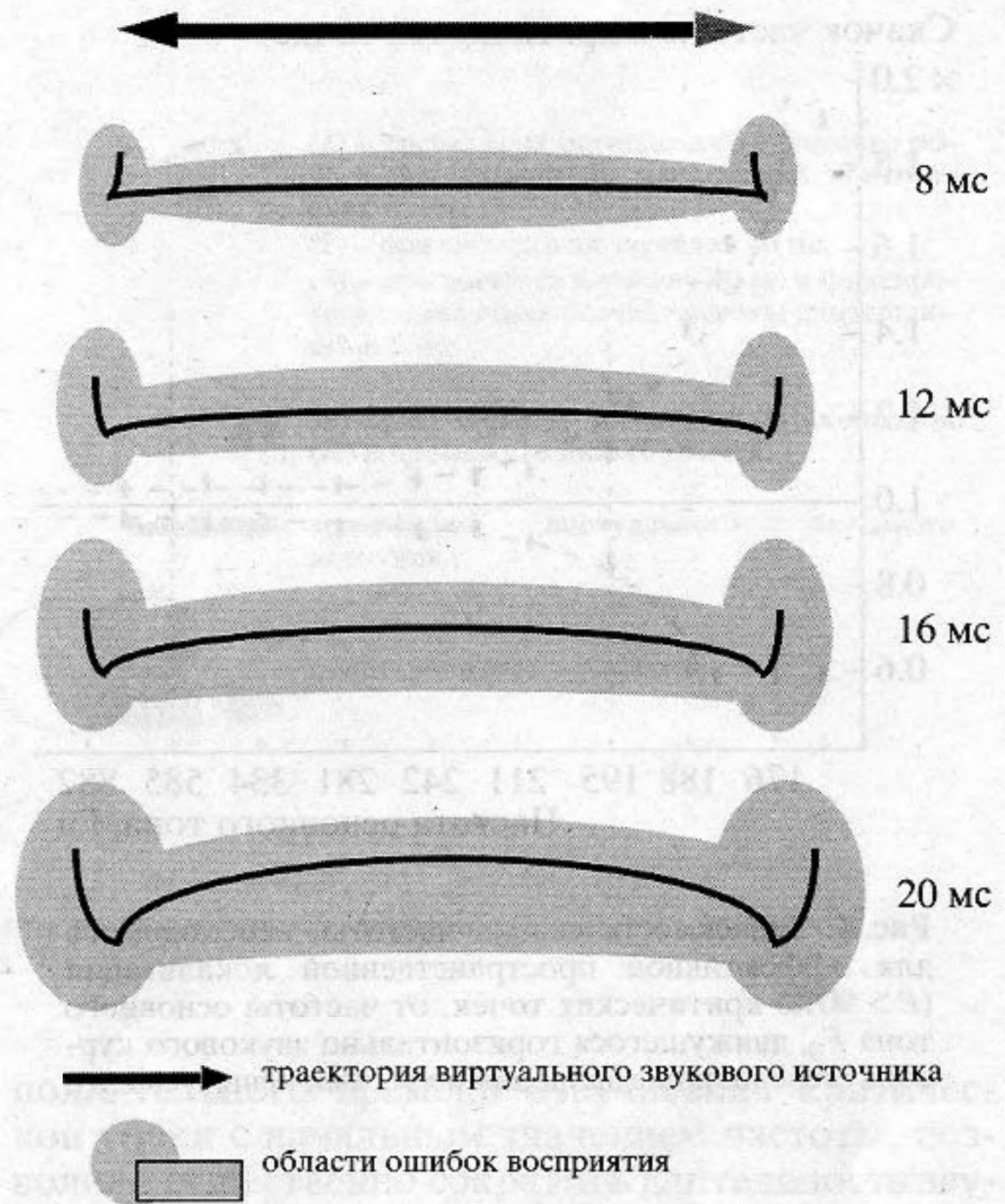


Рис. 3. Субъективные образцы отрезка прямой линии при разной длительности критических точек.

траектории виртуальной акустической плоскости. Практически, необходимо было определить совокупность параметров виртуального источника звука в области критических точек, облегчающих условия сенсорной фиксации соответствующей позиции, но при этом не создающих условий маскирования соседних точек. Таким требованиям удовлетворяли значения параметров, увеличенные в два раза по сравнению с пороговыми. Это позволило уменьшить дисперсию результатов тестирования, связанную с индивидуальными слуховыми особенностями в группе испытуемых. Полученные результаты имеют гораздо большие величины, чем данные экспериментов [7] или аудиометрические исследования дифференциальных порогов различения частоты перекрывающихся тонов и импульсных сигналов (цит. по [10]). Усредненный по результатам тестирования 8 испытуемых график зависимости скачка частоты, необходимого для эффективной пространственной локализации критических точек, от частоты основного тона движущегося горизонтально звукового курсора приведен на рис. 4.

Здесь хорошо видны три области: ниже 200 Гц, от 200 Гц до 500 Гц, и выше 500 Гц частоты основного тона. Соответственно, в первой области эффективное значение скачка частоты в крайних точках составляет более 1.2, в третьей – не более



Рис. 4. Зависимость скачка частоты, необходимого для эффективной пространственной локализации ( $P > 90\%$ ) критических точек, от частоты основного тона  $F_0$ , движущегося горизонтально звукового курсора (усредненные значения по восьми испытуемым).

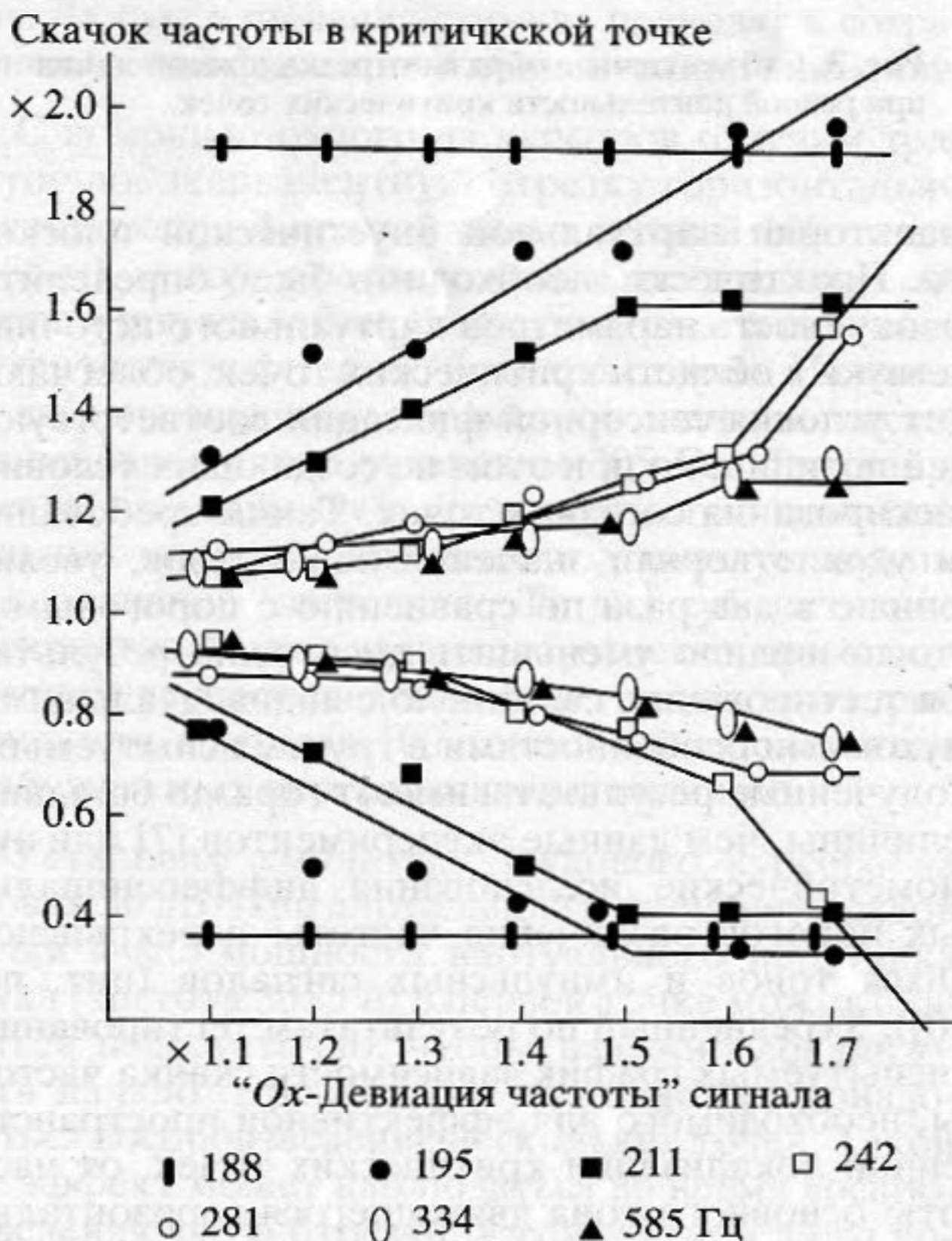


Рис. 5. Зависимость скачка частоты в критических точках, необходимого для эффективной локализации этих точек ( $P > 90\%$ ), от "Ox-Девиации частоты" и частоты основного тона  $F_0$  (параметр кривых), движущегося горизонтально звукового курсора (усредненные значения для одного испытуемого).

1.02, и в средней области – зависит от частоты, изменяясь в пределах 1.02–1.2. Но такие условия были приемлемы в отсутствие девиации основного тона звукового курсора вдоль осей  $Ox$ ,  $Oy$ . В случае разработки произвольных звуковых объектов эти параметры имеют большое значение и изменяются от 1.05 до 1.6. Следовательно, не менее важно было определить зависимость необходимой величины скачка частоты в критической точке с устанавливаемым распределением девиации частоты основного тона звукового курсора в пределах виртуальной акустической плоскости. Исследование проводили для семи значений основного тона ( $F_0$ ): 188, 195, 211, 242, 281, 334 и 585 Гц, девиации частоты по оси  $Ox$  от 1.1 до 1.7 для горизонтально движущегося звукового курсора. Результаты тестирования одного из испытуемых приведены на рис. 5. Практически значимая величина  $D_x$  лежит в интервале от 1.05 до 1.6 и зависит от частоты основного тона. Девиация частоты вдоль оси  $Ox$  линейно транспортирует правую крайнюю точку относительно левой крайней точки вверх на величину, зависящую от  $D_x$ , при этом величина скачка частоты, относительно значения  $F_0 \times D_x$ , для правой крайней точки не изменяется.

Конечно, при более подробном рассмотрении процессов восприятия в критических точках линейного отрезка, в частности при циклическом реверсировании озвучивания, можно найти различия при изменении знака скачка частоты и девиации основного тона вдоль оси  $Ox$ . Но практического значения этот эффект не имеет, так как величина изменения частоты в критической точке существенно превышает значение порога восприятия различия частоты двух соседних точек. Большее значение имеет знак девиации частоты в критической точке, который, вероятно, имеет смысл связать с ординатой виртуальной акустической плоскости, так как и девиация частоты основного тона уже зависит от этого параметра [11, 12]. В результате анализа экспериментальных данных и уточняющих тестов, параметры критических точек для частот основного тона, превышающих 200 Гц (см. рис. 4, 5), были определены следующими условиями:

- виртуальная акустическая плоскость делится на два сегмента в точке с ординатой  $K$ ;
- если значение ординаты критической точки меньше или равно  $K$ , девиация повышает частоту данной точки в  $H$  раз;
- если значение ординаты критической точки больше  $K$ , девиация понижает частоту данной точки в  $L$  раз.

Эмпирические значения указанных величин для основного тона, не превышающего 550 Гц, получились следующие:  $H = 1.2$ ,  $L = 1.2$ ,  $K = 43/128$ , принимая за начало координат левую нижнюю

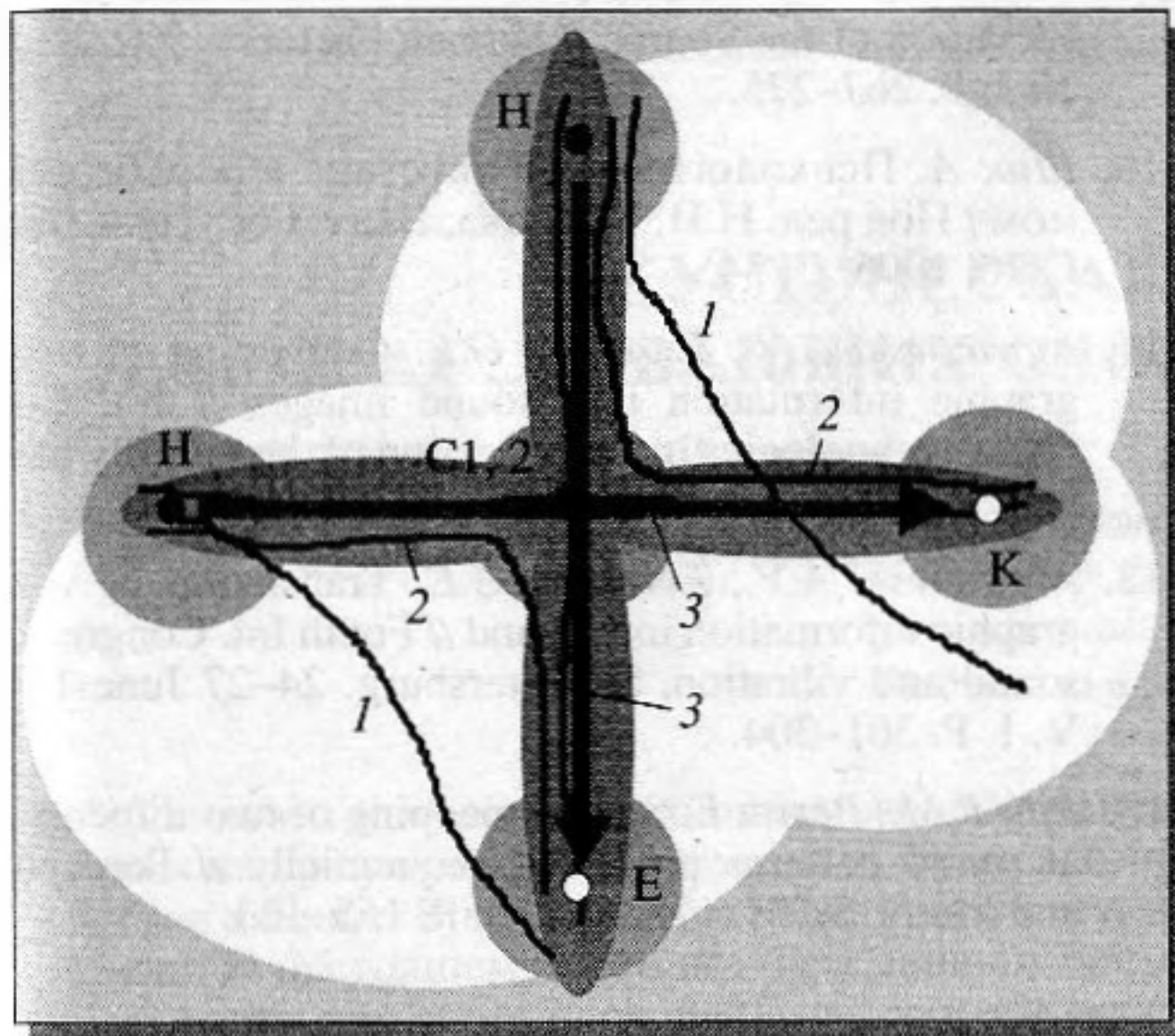


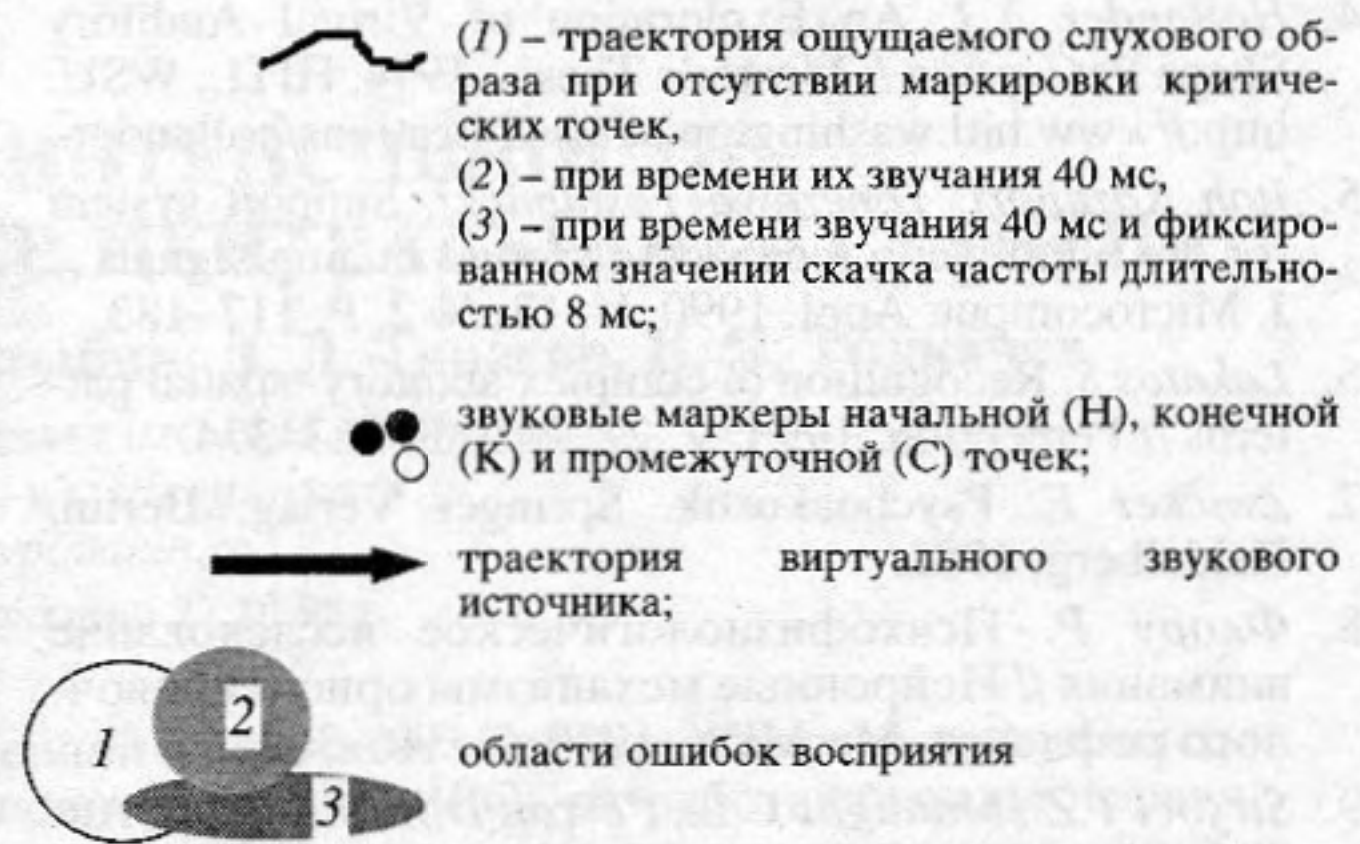
Рис. 6. Восприятие виртуального звукового объекта, сформированного одновременным движением нескольких звуковых источников.

точку виртуальной акустической плоскости. Выявленные закономерности управления параметрами виртуального отрезка прямой линии, расположенного горизонтально, были проверены в аналогичных экспериментах с вертикальным расположением отрезка в акустической плоскости. Дополнительной коррекции или существенных изменений в параметрах управления акустическими курсорами не потребовалось.

Девиация (скачок) частоты критических точек виртуального звукового объекта оказалась эффективным управляемым параметром, и в условиях дефицита времени позволяет удобно маркировать точки, необходимые для переключения внимания пользователя аудиодисплея.

На рис. 6 показаны траектории восприятия виртуального звукового объекта, сформированного одновременным движением двух звуковых источников при отсутствии маркировки критических точек (1), при времени их звучания 40 мс (2), при времени звучания 40 мс и фиксированном значении скачка частоты длительностью 8 мс (3). Хорошо видно сокращение области неопределенной пространственной локализации восприятия траектории движения виртуальных источников. В отличие от известной модели взаимодействия двух звуковых потоков [13], в которой основным фактором слияния или разделения, является временной, в нашей модели введена частотная маркировка критических точек управления вниманием слушателя.

Разделение функций переключения внимания к области, требующей точной пространственной локализации путем скачкообразного изменения частоты и фиксации внимания со помощью до-



полнительного времени озвучивания критической точки с начальным значением частоты, позволило существенно сократить длительность звучания критических точек и стабилизировать восприятие границ виртуального звукового объекта. Предложенный метод позволяет реализовать новые возможности управления пространственными параметрами объектов виртуальных акустических сред.

Работа выполнена по проекту “Пакет прикладных программ преобразования графического изображения в структурированное звуковое поле для слепых пользователей ПЭВМ” (№ гос. регистрации 01.960.010307) в рамках Государственной Научно-технической Программы “ИНФОРМАТИЗАЦИЯ РОССИИ”.

Авторы выражают искреннюю признательность редакции и рецензенту за ценные замечания, внимательное ознакомление и всесторонний анализ, способствующие улучшению содержания работы.

#### СПИСОК ЛИТЕРАТУРЫ

1. Agranovski A.V., Evreinov G.E., Yashkin A.S. Graphic Audio Display for the Blind // Interdisciplinary Aspects in Computers Helping People with Special Needs. 5th Int. Conf. ICCHP'96 Linz, Austria, July 1996. P. 539–542.
2. Евреинов Г.Е., Куркин С.А. Перспективы введения зрительной информации путем адекватной стимуляции слухового анализатора // Акуст. журн. 1996. Т. 42. № 5. С. 629–634.
3. Аграновский А.В., Евреинов Г.Е. Способ формирования акустического образа // Решение о выдаче патента на изобретение Российской Федерации от

- 01.12.97 г. по заявке № 96105192/14 (008800) от 19.03.96 г., МКИ6 А 61 Е 9/08, G 09 В 21/00, RU.
4. *Hollander A.J.* An Exploration of Virtual Auditory Shape Perception // Master's Thesis, 1994. HITL, WSU, <http://www.hitl.washington.edu/publications/hollander>.
  5. *Itoh Kazunori, Yonezawa Yoshimichi.* Support system for the blind using feedback of sound imaging signals // *J. Microcomput. Appl.* 1990. V. 13. № 2. P. 117–183.
  6. *Lakatos S.* Recognition of complex auditory-spatial patterns // *Perception.* 1993. V. 22. № 3. P. 363–374.
  7. *Zwicker E.* *Psychoakustik.* Springer Verlag, Berlin, Heidelberg, 1982.
  8. *Флору Р.* Психофизиологическое исследование внимания // Нейронные механизмы ориентировочного рефлекса. М.: МГУ, 1970. С. 346–360.
  9. *Strybel T.Z., Manligas C.L., Perrott D.R.* Minimum Audible Movement Angle as a Function of the Azimuth and Elevation of the Source // *Human Factors.* 1992. V. 34. № 3. P. 267–275.
  10. *Шук А.* Психологическая акустика в борьбе с шумом / Под ред. Н.И. Иванова; Балт. Гос. Техн. Ун-т, СПб, 1995. С. 34.
  11. *Agranovski A.V., Evreinov G.E.* Converting of visual graphic information into sound images // *Int. Conf. "New technologies in the education of the visually handicapped"*. Paris, 10–11 June 1996. V. 2. P. 28–30.
  12. *Arganovski A.V., Evreinov G.E.* Translating of visual graphic information into sound // *Fouth Int. Congress on sound and vibration.* St. Petersburg, 24–27 June 1996. V. 1. P. 301–304.
  13. *Ruff R.M., Perret E.* Spatial mapping of two-dimensional sound patterns presented sequentially // *Perceptual and Motor Skills.* 1982. V. 55. P. 155–163.

## Creating Virtual Sound Objects

A. V. Agranovskii and G. E. Evreinov

How is an auditory contour pattern perceived? What additional attributes must be adjusted to stabilize the perception of the boundaries of a virtual sound pattern in order to form an auditory perception isomorphous to the visual one? A possibility to correct the perception of subjective patterns through controlling the parameters (the frequency deviation and the duration at the critical points) of virtual sound patterns created by the illusion of the apparent motion of acoustic sources is discussed.