PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 : G10H 1/02

A1

(11) International Publication Number: WO 91/10987

(43) International Publication Date: 25 July 1991 (25.07.91)

(21) International Application Number: PCT/US91/00223

(22) International Filing Date: 17 January 1991 (17.01.91)

(30) Priority data:
465,732 18 January 1990 (18.01.90) US

(81) Designated States: DE, GB, JP.

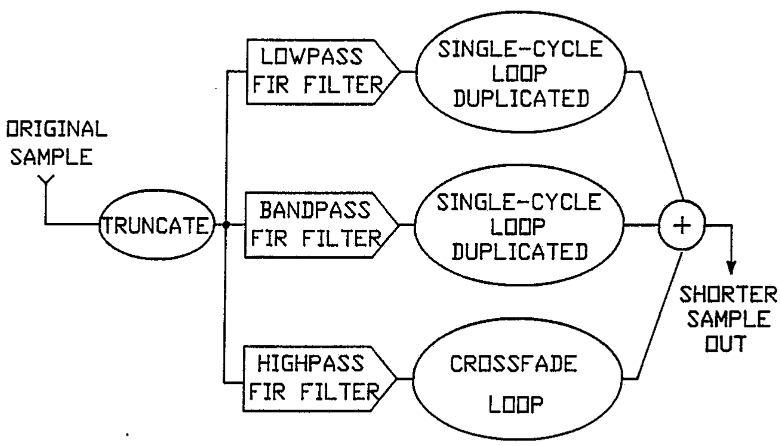
Published
With international search report.

(71) Applicant: E-MU SYSTEMS, INC. [US/US]; 1600 Green Hills Road, Scotts Valley, CA 95066 (US).

(72) Inventors: MONAHAN, Kevin, J. ; 417 Broadway, Santa Cruz, CA 95060 (US). MURRAY, Donna, L. ; 1420 Eastcrest Lane, Santa Cruz, CA 95062 (US).

(74) Agent: BALDWIN, Stephen, E.; Flehr, Hohbach, Test, Albritton & Herbert, Four Embarcadero Center, Suite 3400, San Francisco, CA 94111-4187 (US).

(54) Title: DATA COMPRESSION OF SOUND DATA



ORIGINAL SAMPLE IS TRUNCATED, BANDSPLIT (IN THIS CASE, INTO 3 BANDS), SEPARATELY LOOPED, AND THEN RECOMBINED. RESULT IS A MUCH SHORTER SAMPLE. ALL BANDSPLIT FILTERS MUST BE OF THE SAME ORDER TO INSURE PHASE CONSISTENCY UPON RECOMBINATION.

(57) Abstract

A data compression method and apparatus (Fig. 14) for the compression of sound data utilized in digital sampling keyboard instruments. The present invention reduces memory requirements for sampled sounds without compromising sound quality. Samples have an attack portion and a cross/faded loop portion. Sound data between the attack portion and just before the loop start portion are deleted. The remaining attack and loop portions are spliced to form a spliced data sample (Fig. 6).

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	**ES**	Spain	**MG**	Madagascar
AU	Australia	**FI**	Finland	**ML**	Mali
BB	Barbados	**FR**	France	**MN**	Mongolia
BE	Belgium	**GA**	Gabon	**MR**	Mauritania
BF	Burkina Faso	**GB**	United Kingdom	**MW**	Malawi
BG	Bulgaria	**GN**	Guinea	**NL**	Netherlands
BJ	Benin	**GR**	Greece	**NO**	Norway
BR	Brazil	**HU**	Hungary	**PL**	Poland
CA	Canada	**IT**	Italy	**RO**	Romania
CF	Central African Republic	**JP**	Japan	**SD**	Sudan
CG	Congo	**KP**	Democratic People's Republic of Korea	**SE**	Sweden
CH	Switzerland			**SN**	Senegal
CI	Côte d'Ivoire	**KR**	Republic of Korea	**SU**	Soviet Union
CM	Cameroon	**LI**	Liechtenstein	**TD**	Chad
CS	Czechoslovakia	**LK**	Sri Lanka	**TG**	Togo
DE	Germany	**LU**	Luxembourg	**US**	United States of America
DK	Denmark	**MC**	Monaco		

DATA COMPRESSION OF SOUND DATA

Background of the Invention

The present invention relates to data compression of sound data, and more particularly to the data compression of sound data utilized in digital sampling keyboard instruments.

Since the introduction of digital sampling keyboard instruments, the desire to compress sound data into smaller memories without compromising sound quality is ever increasing. In recent years, limiting bit resolution (8 to 12 bits) and sample rates (less than 44.1 Khz) have been two common methods to reduce memory size. But since the introduction of the compact disk (CD), resolution less than 16-bit and 44.1 Khz has largely been deemed unacceptable.

Another common approach, looping, involves repeating a section of data during the time a key is depressed. Two common types of loops are single period forwards loops and cross-faded forwards loops (see Figs. 1 and 2). Single period (or single cycle) loops characteristically sound quite static, as only one period is repeated. They work best on solo instruments with non-complex harmonic structures. Longer loops, on the other hand, are required for ensemble sounds and harmonically complex solo sounds. Often, the sound data must be processed to avoid pops in the loop.

This process is called cross/fade looping. Portions of the sound at the loop start and end points are faded in and out of the loop. Obviously, the longer, cross-faded loop contains more dynamics than a single-cycle loop. However, some lower frequency phase cancellation occurs as a result.

The start point of a cross/faded loop must begin after the attack phase of the sound has passed and the sound becomes more stable. The problem here is that it often takes a while for a sound to become stable. If a loop is started too close to the attack, poor loops result due to large fluctuations in phase and amplitude, and there is a high risk of attack data becoming part of the loop.

Yet another method for reducing memory is to simply take fewer samples of a given instrument across the keyboard. A single sample of a violin will use less memory than one that has been sampled every half octave. The problem here is that the realism of the sound disintegrates rapidly when too few samples are used to represent a fixed formant instrument.

Summary of the Invention

It is an object of the present invention to provide an improved data compression method and apparatus to be utilized with digital sampling keyboard instruments.

A more particular object of the invention is to reduce memory requirements for sampled sounds without compromising sound quality, using three techniques. Additionally, the third technique improves the defect of formant distortion when sampled sounds are transposed.

Briefly, the present invention is directed toward, in one preferred embodiment, an improved method for processing sound data samples where the data samples have an attack portion and a cross/faded loop portion, including the step of deleting the sound data between the attack portion and just before the loop start portion. The improved method further includes the step of digitally splicing the remaining attack and loop portions to form a spliced data sample.

Additional objects, advantages and novel features of the present invention will be set forth in part in the description which follows and in part become apparent to those skilled in the art upon examination of the following, or may be learned by practice of the invention. The objects and advantages of the present invention may be realized and attained by means of the instrumentalities and combinations which are pointed out in the appended claims.

Brief Description of the Drawings

The accompanying drawings which are incorporated in and form a part of this specification illustrate an embodiment of the inventin and, together with the description, serve to explain the principles of the invention.

Fig. 1 depicts a single-cycle forwards loop.

Fig. 2 depicts a cross-fade looping process.

Fig. 3 depicts a conventional cross-faded loop.

Fig. 4 depicts a conventional cross-faded loop too close to attack.

Fig. 5 depicts cut, copy and paste procedures.

Fig. 6 depicts attack/loop splice.

Fig. 7 depicts piano sample with conventional cross-fade loop.

Fig. 8 depicts piano sample with conventional cross-fade loop closer to attack, showing fluctuation in loop.

Fig. 9 depicts piano sample band split and looped separately.

Fig. 10 depicts piano sample band split and loops equalized.

Fig. 11 depicts piano sample bands recombined with a resultant loop closer to attack.

Fig. 12 depicts formant shifting.

Fig. 13 depicts a diagram of a digital finite impulse response (FIR) filter.

Fig. 14 depicts a diagram of a data compression technique in which lowpass, bandpass and highpass FIR filters are utilized.

Detailed Description of the Invention

Reference will now be made in detail to the preferred embodiment of the invention, an example of which is illustrated in the accompanying drawings. While the invention will be described in conjunction with the preferred embodiment, it will be understood that it is not intended to limit the invention to that embodiment. On the contrary, it is intended to cover alternatives, modifications and equivalents as may be included within the spirit and scope of the invention as defined by the appended claims.

One technique according to the present invention for data reduction (Fig. 3-6) utilizes cut and paste editing tools to reduce a sound to its most essential components, attack and loop (sustain). Fig. 3 shows a string sample that has been cross-fade looped well after the attack. Sonically, this example is correct, but requires more memory than is desirable (57K). In Fig. 4, the same sample has been looped much closer to the attack of the sound, producing the desired memory reduction (22K), but now the loop contains elements of the attack. Because of the instability of the sound at that point in time, the loop has an undesirable amount of fluctuation.

Returning to Fig. 3, it can be seen that the sound data between the attack (approximately 125 ms) and just before the loop start (approximately 100 ms) can be deleted. The remaining portion (the attack and loop) can be digitally spliced together with up to 100 ms X/fade time. The X/fade will prevent any audible pop in the splice and the fade time is limited by the size of data before the loop start, which in this case is 100 ms or 4410 bytes at 44.1 Khz sample rate. (See Fig. 5)

The resulting sample (Fig. 6) not only saves memory but can sound significantly better than the example in Fig. 3, because the unstable portion of the sample has been eliminated.

A second technique (Figs. 7-11) according to the present invention utilizes a phase linear filter to separate a sample into multiple bands that can be individually processed and looped much closer to the attack, then digitally recombined. The use of finite impulse response digital filters of consistent order between bands insures no phase distortion of the result after recombining.

Fig. 7 shows a piano sample that has been crossfade looped. A shorter sample is desired. In this example, a single cycle loop of the original sample would sound very static and unnatural. Use of a cross-faded loop closer to the attack results in excessive tremolo effects due to the amound of animation still present in the looping area of the sound and the phase cancellation byproducts of crossfading, as shown in Fig. 8.

The variations of lower frequency components in the loop are what cause the the undesirable tremolo effects, but variations of higher frequency components in the loop are useful to maintain an animated sound. Bandsplitting the shorter sample allows the low-frequency components to be single-cycle looped, and the high-frequency components to use cross-faded loops. The result after recombining the bands is a loop that sounds stable but still animated.

In Fig. 9, the piano sample has been split into three bands using a lowpass, bandpass and highpass phase-linear filter. Band A has been lowpassed, leaving mostly the fundamental frequency (51 hz, G#0). Band B has been bandpassed, leaving only the second harmonic. Band C has been highpassed, leaving the remainder of the sound.

Band A is looped using a single cycle loop and band B is looped at the same length (which is actually a double cycle loop). Band C is looped using a much longer crossfade loop. The only restriction here is that the longest loop length of all the bands must be an integer multiple of the other loop lengths to allow for proper recombination later. In this case, loops in A and B are 850 bytes and the loop in C is 45900 bytes (54 times as long).

In order to recombine the three bands back into one sample, the loop lengths must first be equalized (Fig. 10). This is accomplished by first copying the loop data in band A many times until a loop length equal to that of C is achieved. In this case, multiplying the loop 54 times provides the correct loop length, but the loop start must also occur at exactly the same point. Simply moving the loop start points of band A to that of band C may result in a less desirable band A loop. Therefore, the loop data in band A should be copied an additional number of times until enough data is created to produce a loop length of 45900 bytes at a start point equal to band C, 94779 bytes. This process is repeated for band B, yielding three bands that all have loops which start at 94779 bytes and are 45900 bytes in length.

With the loops equalized, the three bands can now be recombined (Fig. 11). The resultant sample has a very natural sustain with some motion in the higher frequencies and very little in the lower ranges. If the original sample had been looped using conventional X/fade looping methods, it would be necessary to start the loop much further from the attach to achieve a similar loop stability (Fig. 7). Otherwise, the sample would contain phase cancellation defects in the low end, which can be observed in Fig. 8.

A third data compression technique according to the present invention combines two or more pitches of ensemble sounds into one sample, thereby creating larger sounds in less memory, as well as reducing formant distortion due to pitch-shifting.

When a fixed-formant sampled sound is shifted flat or sharp upon playback, it sounds uncharacteristic. An obvious example would be a single section vocal "aah" sample stretched up and down an octave. The sizes of the vocalists seem to grow and shrink unrealistically. This phenomenon is sometimes called "munchkinization."

Fig. 12 illustrates formant transposition as a result of pitch-shifting the vowel "ah" from A 440 Hz to F 349 Hz and from F 349 Hz to A440 Hz. When compared with the original pitches, the transposed versions exhibit a deviation in formant location.

By digitally re-tuning F 349 Hz to A440 Hz, then digitally combining it with the original A 440 Hz sample, the resultant formant location more closely approximates that of the original A 440 Hz sample. Also, since the combined version contains formant characteristics of both pitches, the effective transposition range has been increased and a larger section sound produced within each sample.

Referring now to Fig. 13, there is shown therein a digital finite impulse response filter. The filter coefficients, C_i, must all be real to insure a linear phase response. The order of the filter is the number of stages, N.

Fig. 14 illustrates a data compression technique according to the present invention in which the original sample is truncated, band split (in this case, into three bands), separately looped, and then recombined. The result is much shorter sample. All band split filters in Fig. 14 are of the same order to insure phase consistency upon recombination.

In Fig. 14, after truncation lowpass, bandpass and highpass filtering is performed, as described above. The output of the lowpass FIR filter is then single cycle loop duplicated.

The output of the bandpass FIR filter is single cycle loop duplicated.

The output of the highpass FIR filter is cross/fade looped. The looped bands are then combined, as shown in Fig. 14.

The aspects of the present invention can be achieved by utilizing suitable digital sampling keyboard instruments such as the EMULATOR III which is manufactured by the same applicant as the present invention herein, namely E-mu Systems, Inc. of Scotts Valley, CA. Also, commercially available sound processing software can be utilized in conjunction with such a suitable digital sampling instrument to provide data compression of sound data according to the present invention.

The foregoing description of the preferred embodiment of the present invention has been presented for purposes of illustration and description. It is not intended to be exhaustive or to limit the invention to the precise form disclosed, and many modifications and variations are possible in light of the above teaching. The preferred embodiment was chosen and described in order to best explain the principles of the invention and its practical applications to thereby enable othrs skilled in the art to best utilize the invention and with various embodiments and with various modifications as are suited to the particular use contemplated. It is intended that the scope of the invention be defined only by the claims appended hereto.

What is Claimed is:

1. In a data compression method for the compression of sound data, the method comprising the steps of

processing a sound data sample having an attack portion and a cross/faded loop portion, including the step of deleting the sound data between the attack portion and just before the loop start portion, and

digitally splicing the remaining attack and loop portions to form a spliced data sample.

2. The method as in Claim 1 including the step of digitally splicing the remaining attack and loop portions with a predetermined cross fade time.

3. The method as in Claim 2 wherein the cross fade time is approximately 100 milliseconds.

4. Data compression apparatus for the compression of sound data, the apparatus comprising

means for processing a sound data sample having an attack portion and a cross/faded loop portion, including means for deleting the sound data between the attack portion and just before the loop start portion, and

means for digitally splicing the remaining attack and loop portions to form a spliced data sample.

5. In a data compression method for the compression of sound data samples, the method comprising the steps of

splitting a sound data sample into a lowpass band, a bandpass band and a highpass band, such that the lowpass band includes the fundamental frequency of said data sample, the bandpass band includes the second harmonic of said data sample, and the highpass band includes the remainder of the sound, looping the lowpass band and highpass band, using a single cycle loop and a double cycle loop, respectively,

looping the highpass band using a cross fade loop such that the longest loop length is an integer multiple of the other loop lengths, and

recombining the looped bands into a recombined data sample.

6. The method as in Claim 5 wherein the recombining step includes equalizing the respective loop lengths.

7. The method as in Claim 6 wherein the respective loops stars occur at the same respective point.

8. In a data compression method for the compression of sound data sample, the method comprising the steps of

splitting a sound data sample into at least a first band and a second band, such that said first band includes the fundamental frequency of said data sample, and the second band includes the remainder of the sound,

looping the first band using a single cycle loop,

looping the second band using a crossfade loop such as the longest loop length is an integer multiple of the other loop length, and

recombining the loop bands into a recombined data sample.

9. Data compression apparatus for the compression of sound data samples, the apparatus comprising

means for splitting a sound data sample into a lowpass band, a bandpass band and a highpass band, such that the lowpass band includes the fundamental frequency of said data sample, the bandpass band includes the second harmonic of said data sample, and the highpass band includes the remainder of the sound, means for looping the lowpass band and highpass band, using a single cycle loop and a double cycle loop, respectively,

means for looping the highpass band using a crossfade loop such that the longest loop length is an integer multiple of the other loop lengths, and

means for recombining the looped bands into a recombined data sample.

10. In a data compression method for the compression of sound data, the method comprising the steps of

sampling first and second sound data samples at first and second sampling frequencies, respectively,

pitch shifting said data samples to form a formant transposition of said second and first sampling frequencies, respectively,

digitally combining the first and second samples and the transposed samples to form a combined sound data sample.

11. Data compression apparatus for the compression of sound data, the apparatus comprising

means for sampling first and second sound data samples at first and second sampling frequencies, respectively,

means for pitch shifting said data samples to form a formant transpositiion of said second and first sampling frequencies, respectively,

means for digitally combining the first and second samples and the transposed samples to form a combined sound data sample.

12. In a data compression method for the compression of sound data, the method comprising the steps of

sampling first and second sound data samples at first and second sampling frequencies, respectively,

pitch shifting one of said data samples to the other of said data samples to form a formant transposition thereof,

digitally combining the first and transposed samples to form a combined sound data sample.

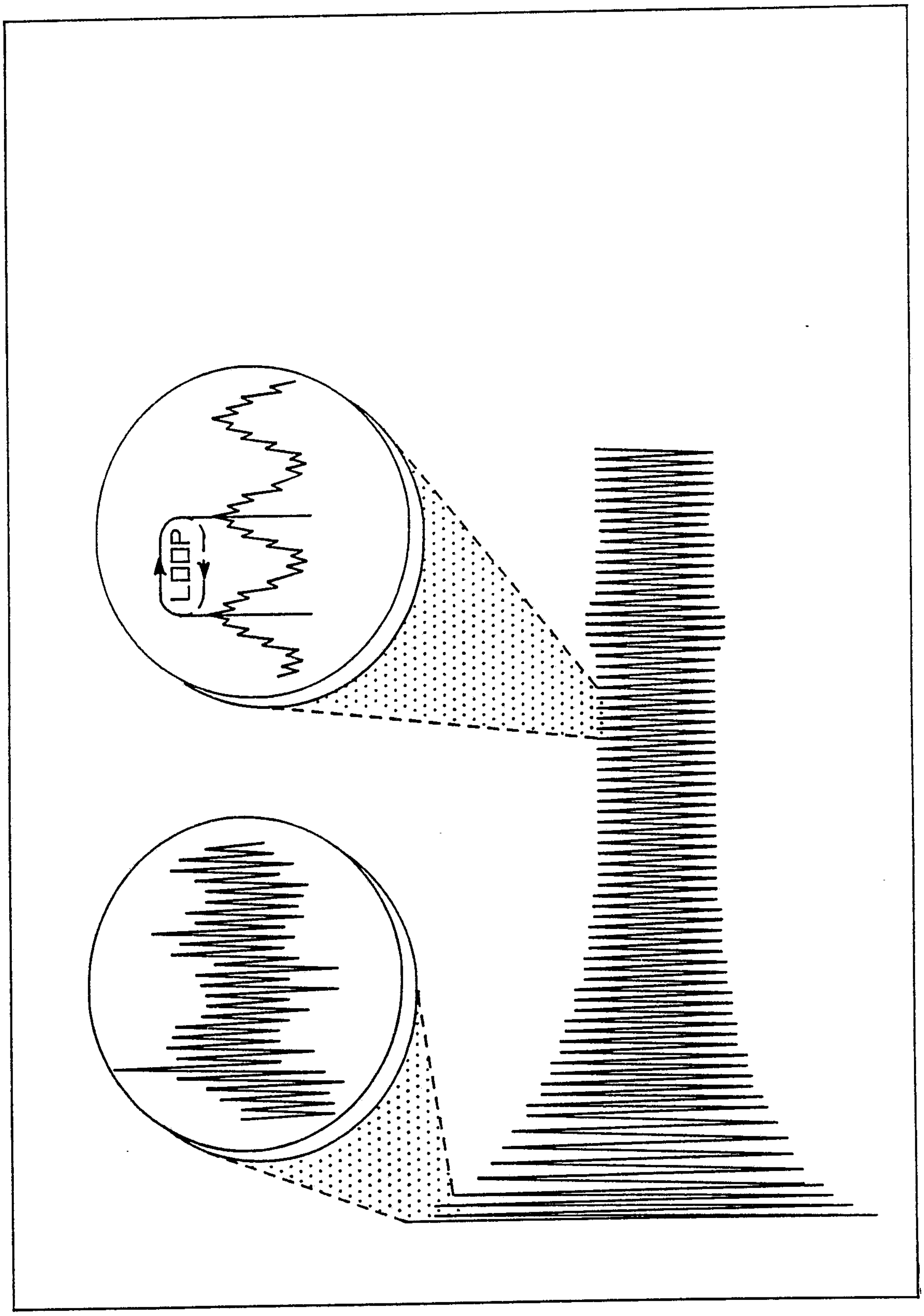
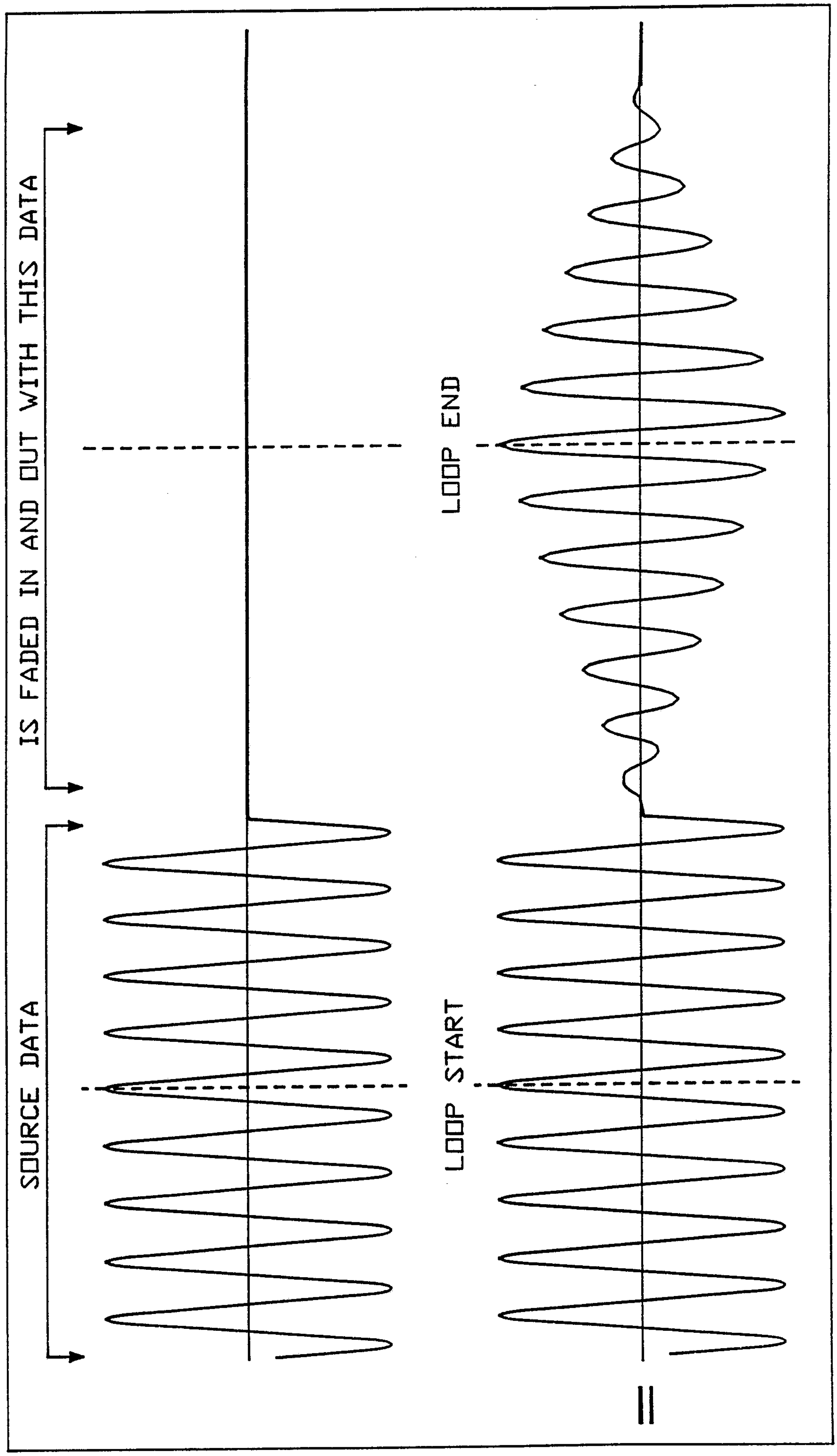


FIG.—1



SOURCE DATA
IS FADED IN AND OUT WITH THIS DATA
LOOP START
LOOP END
||

FIG.-2

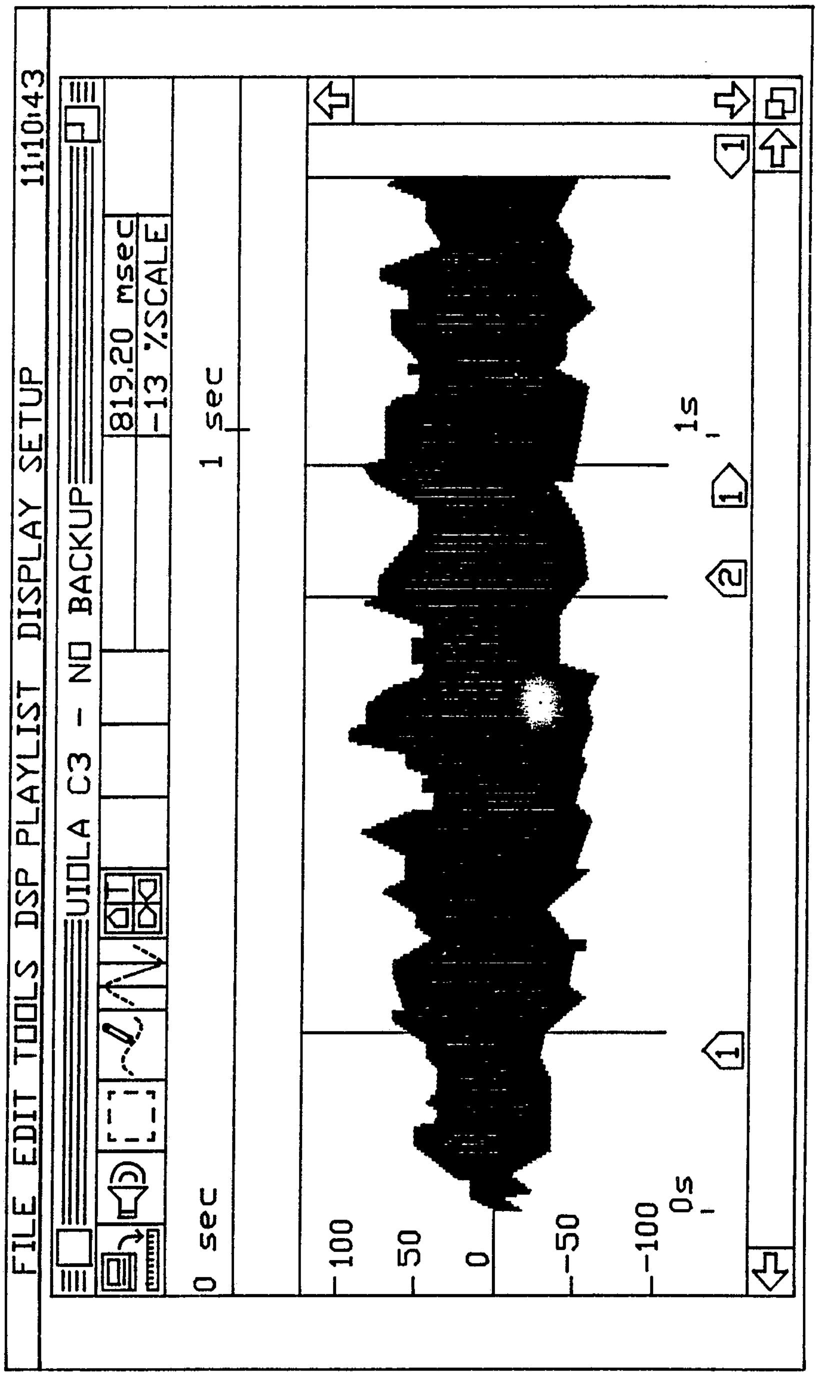

FILE EDIT TOOLS DSP PLAYLIST DISPLAY SETUP 11:10:43
VIOLA C3 - NO BACKUP
819.20 msec
-13 %SCALE
0 sec
1 sec
100
50
0
-50
-100
0s
1s
1
2
1
1

FIG. — 3

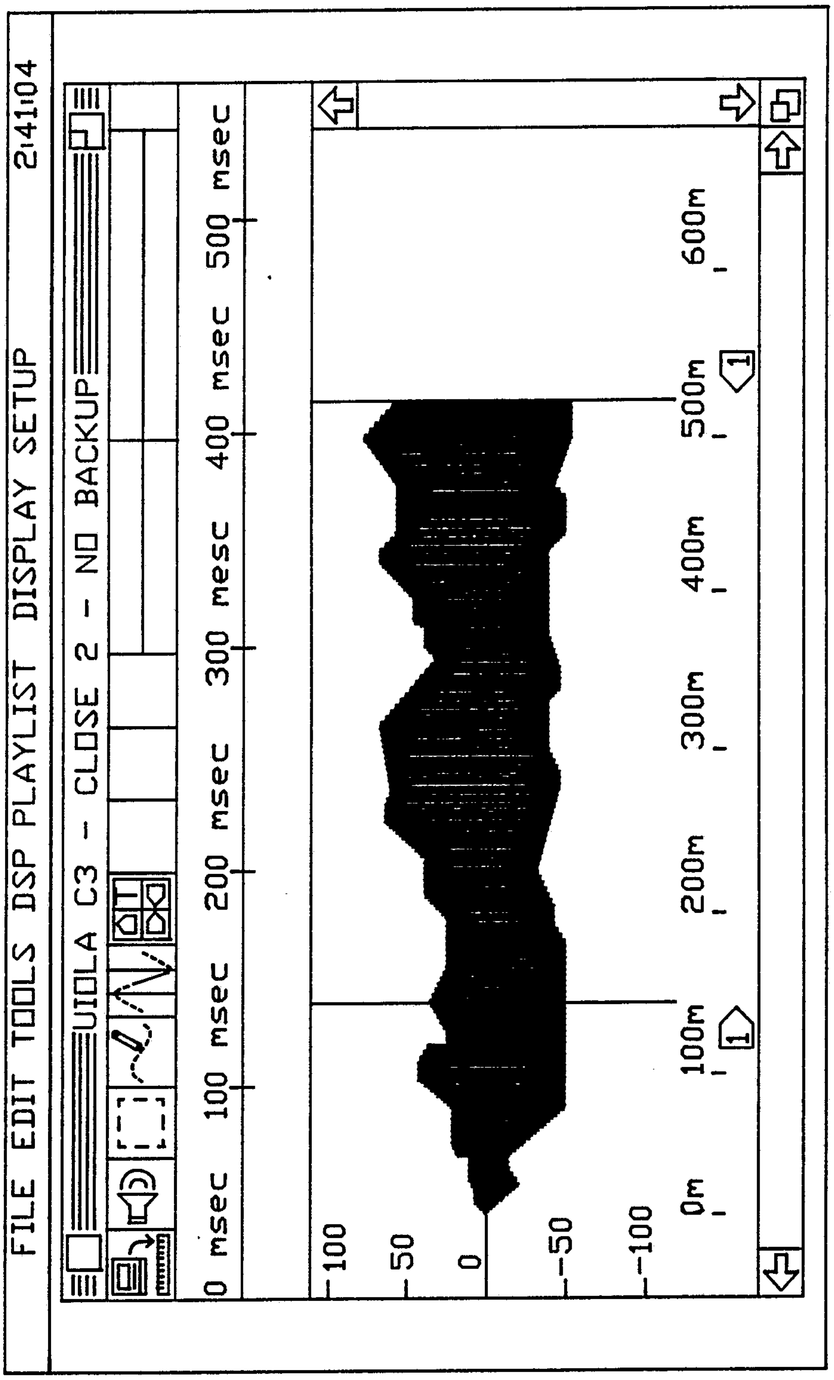

FILE EDIT TOOLS DSP PLAYLIST DISPLAY SETUP
2:41:04
VIOLA C3 - CLOSE 2 - NO BACKUP
0 msec
100 msec
200 msec
300 msec
400 msec
500 msec
100
50
0
-50
-100
0m
100m
200m
300m
400m
500m
600m
1
1

FIG.—4

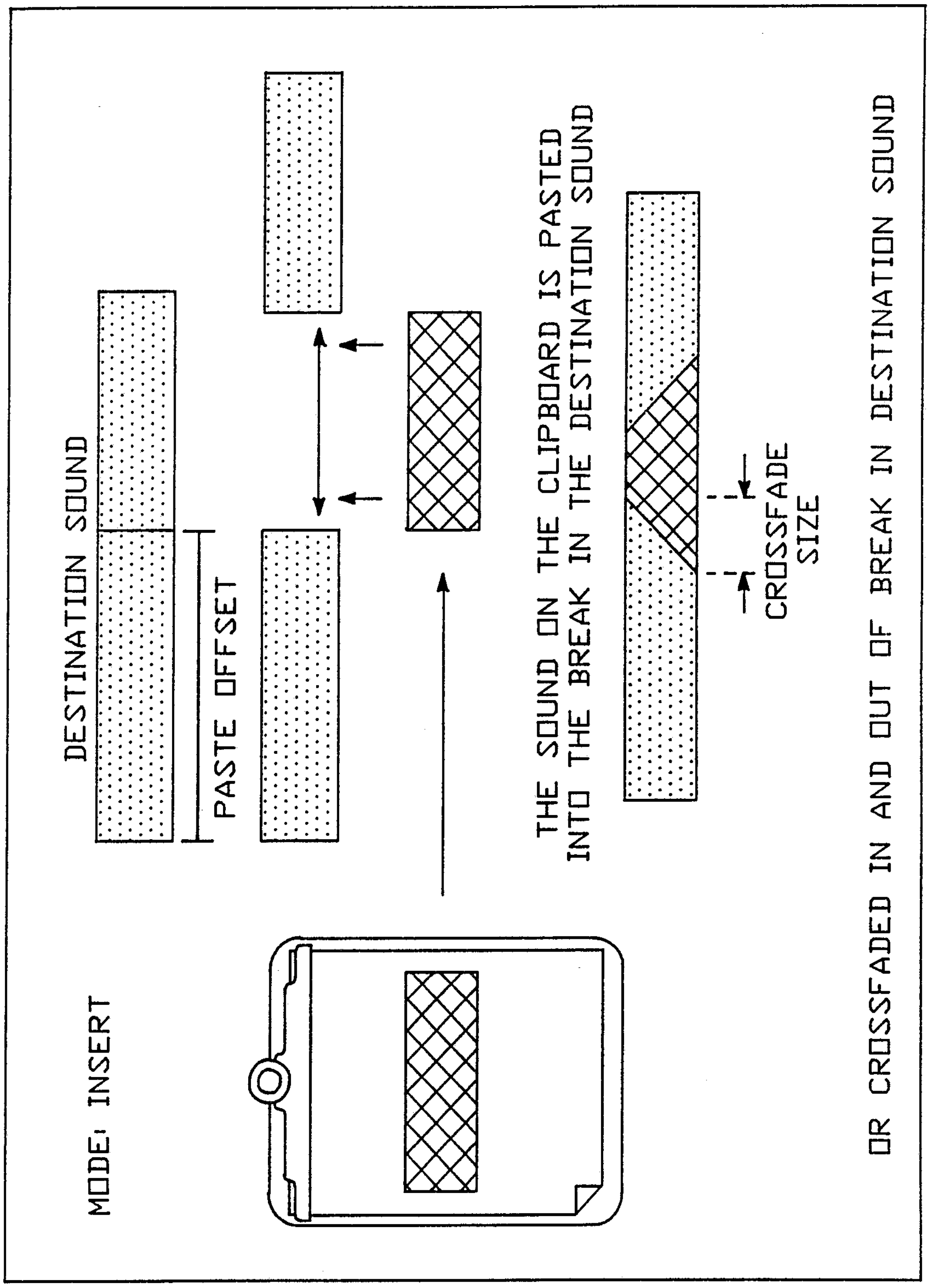

MODE: INSERT
DESTINATION SOUND
PASTE OFFSET
THE SOUND ON THE CLIPBOARD IS PASTED
INTO THE BREAK IN THE DESTINATION SOUND
CROSSFADE
SIZE
OR CROSSFADED IN AND OUT OF BREAK IN DESTINATION SOUND

FIG.—5

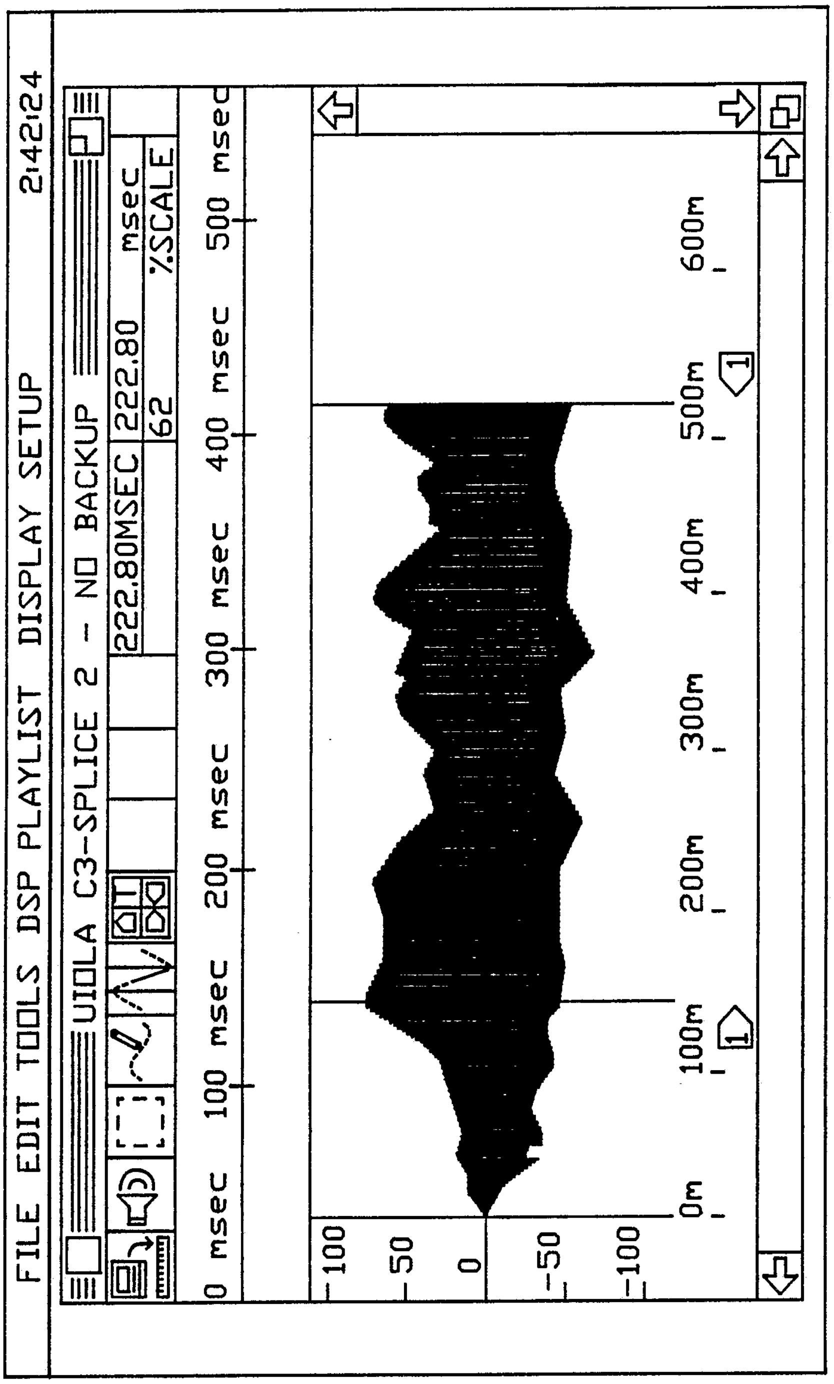

FILE EDIT TOOLS DSP PLAYLIST DISPLAY SETUP
2:42:24
VIOLA C3-SPLICE 2 - NO BACKUP
222.80MSEC
222.80
msec
62
%SCALE
0 msec
100 msec
200 msec
300 msec
400 msec
500 msec
100
50
0
-50
-100
0m
100m
200m
300m
400m
500m
600m
1
1

FIG.—6

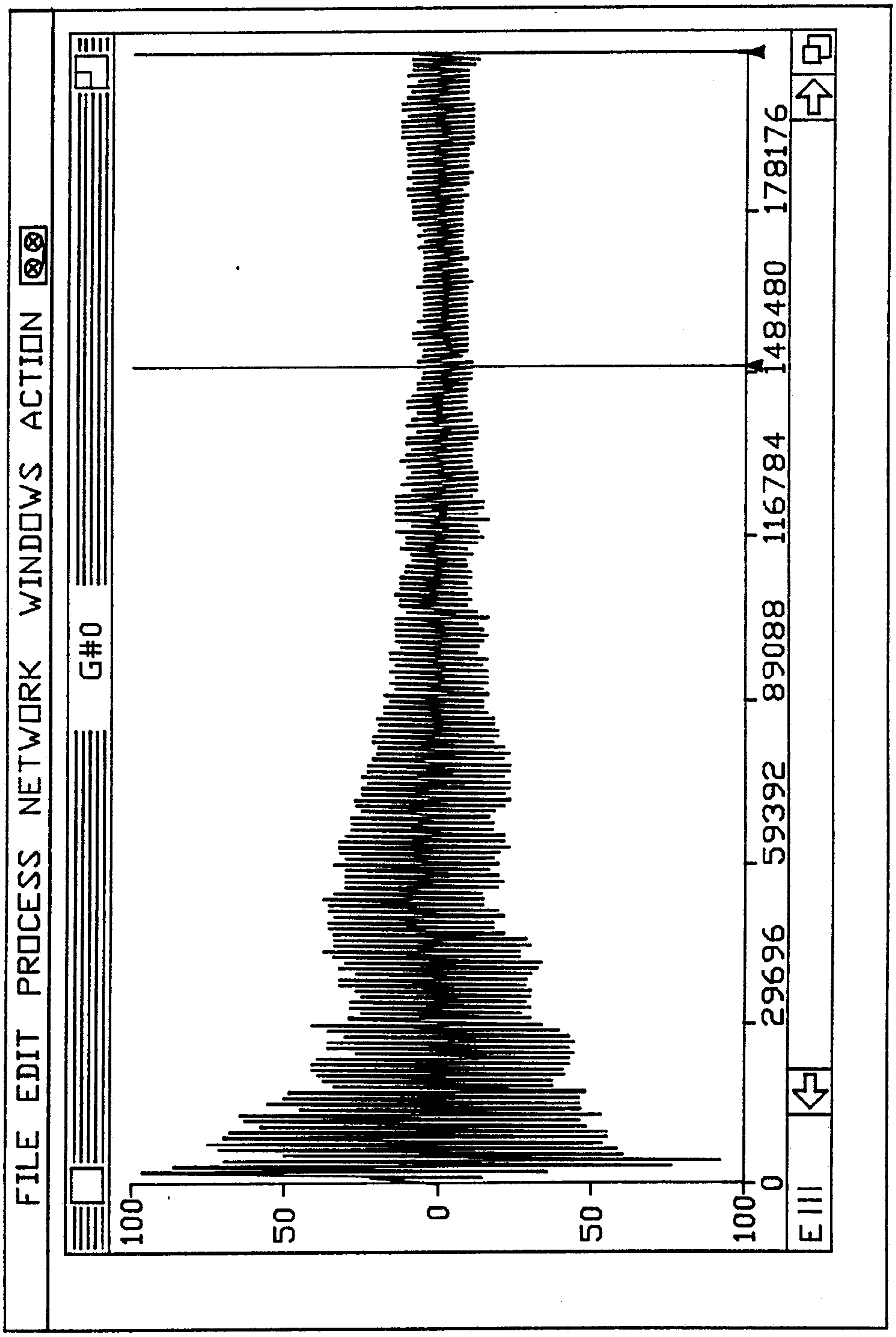

FILE EDIT PROCESS NETWORK WINDOWS ACTION
G#0
100
50
0
50
100
0
29696
59392
89088
116784
148480
178176

FIG.—7

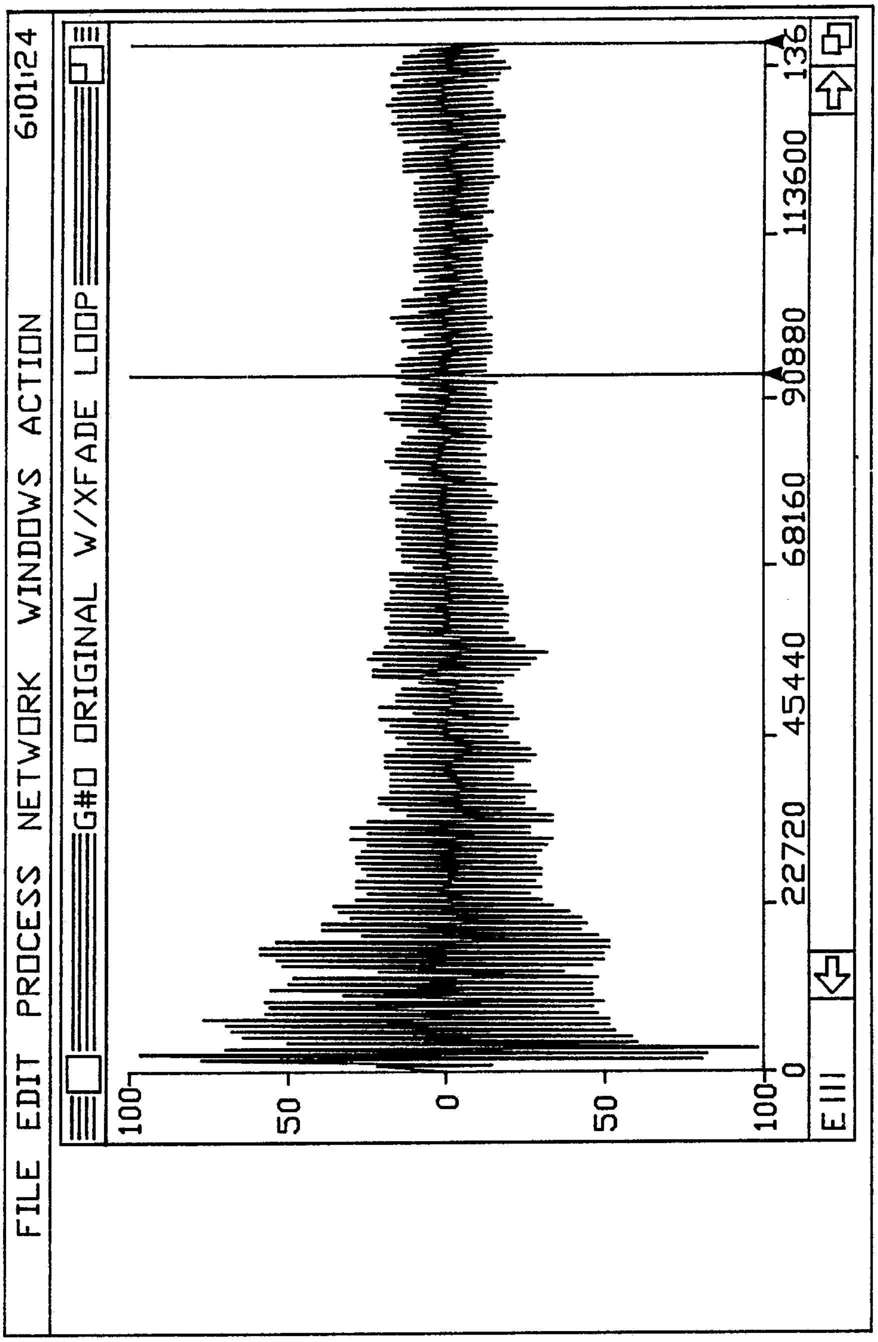

FILE EDIT PROCESS NETWORK WINDOWS ACTION
6:01:24
G#0 ORIGINAL W/XFADE LOOP
100
50
0
50
100
0
22720
45440
68160
90880
113600
136

FIG.—8

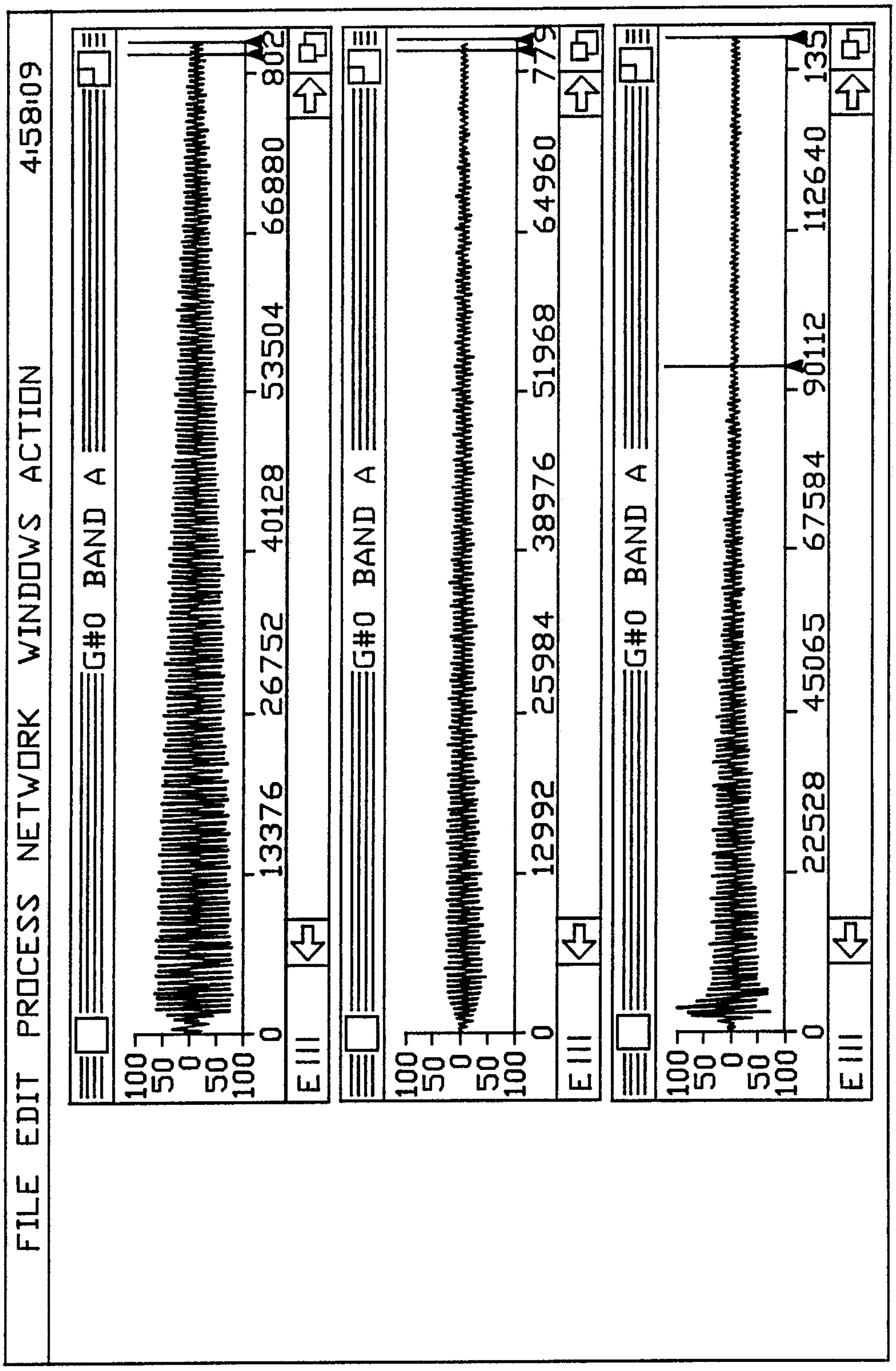

FILE EDIT PROCESS NETWORK WINDOWS ACTION 4:58:09
G#0 BAND A
100 50 0 50 100
0 13376 26752 40128 53504 66880 802
G#0 BAND A
100 50 0 50 100
0 12992 25984 38976 51968 64960 779
G#0 BAND A
100 50 0 50 100
0 22528 45065 67584 90112 112640 135

FIG. — 9

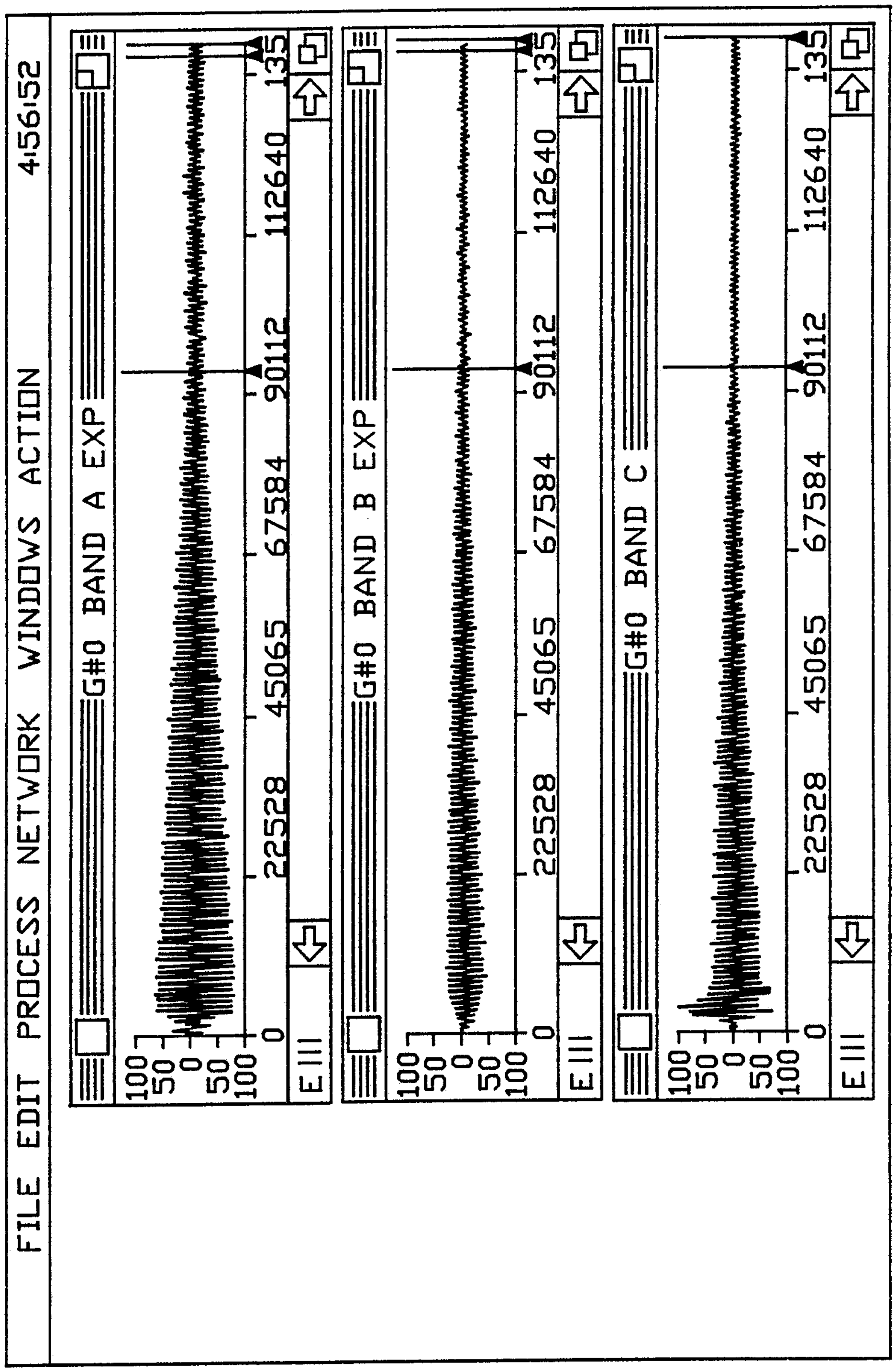

FILE EDIT PROCESS NETWORK WINDOWS ACTION 4:56:52
G#0 BAND A EXP
G#0 BAND B EXP
G#0 BAND C
100 50 0 50 100
0 22528 45065 67584 90112 112640 135

FIG. —10

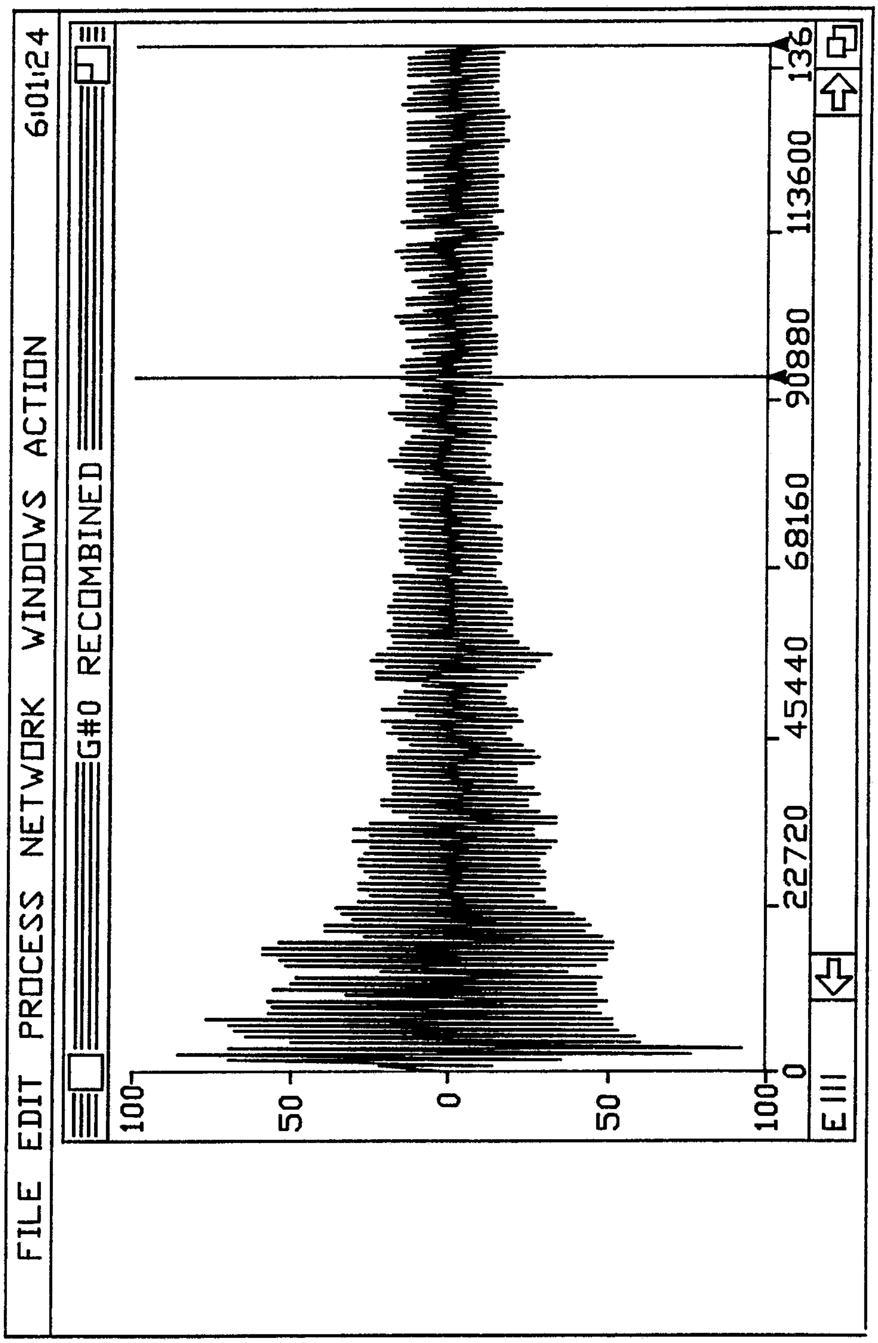

FILE EDIT PROCESS NETWORK WINDOWS ACTION
6:01:24
G#0 RECOMBINED
100
50
0
50
100
0
22720
45440
68160
90880
113600
136

FIG. – 11

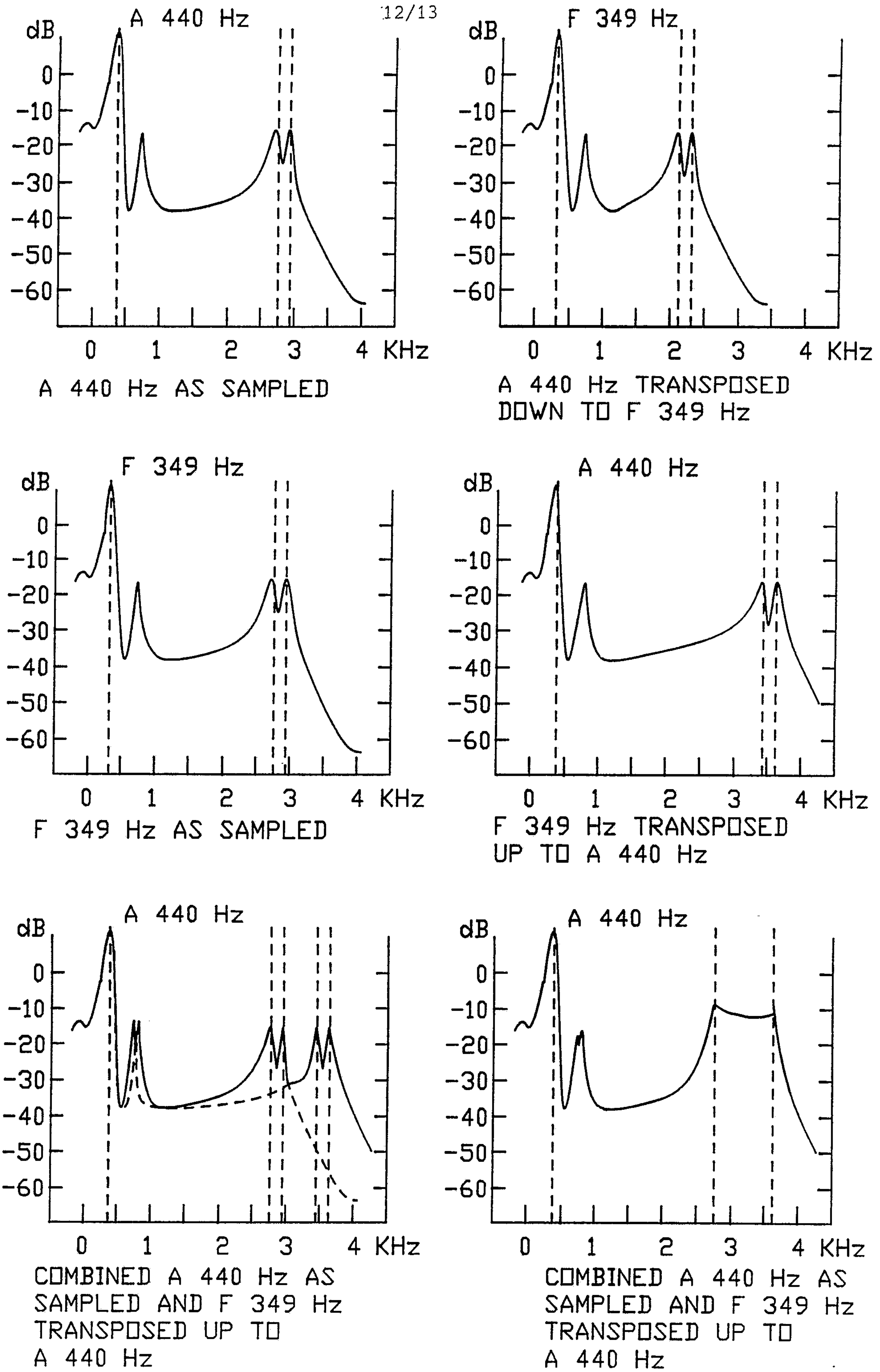
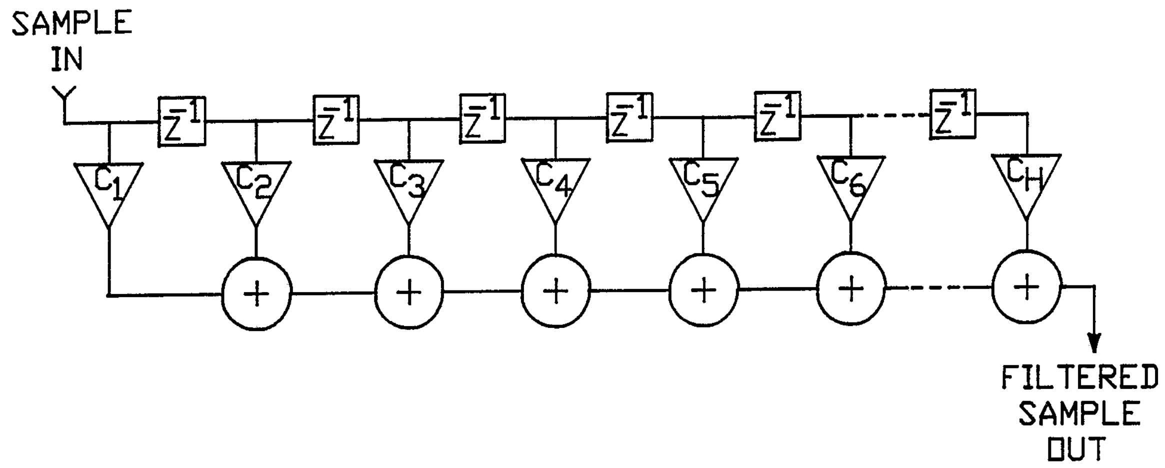
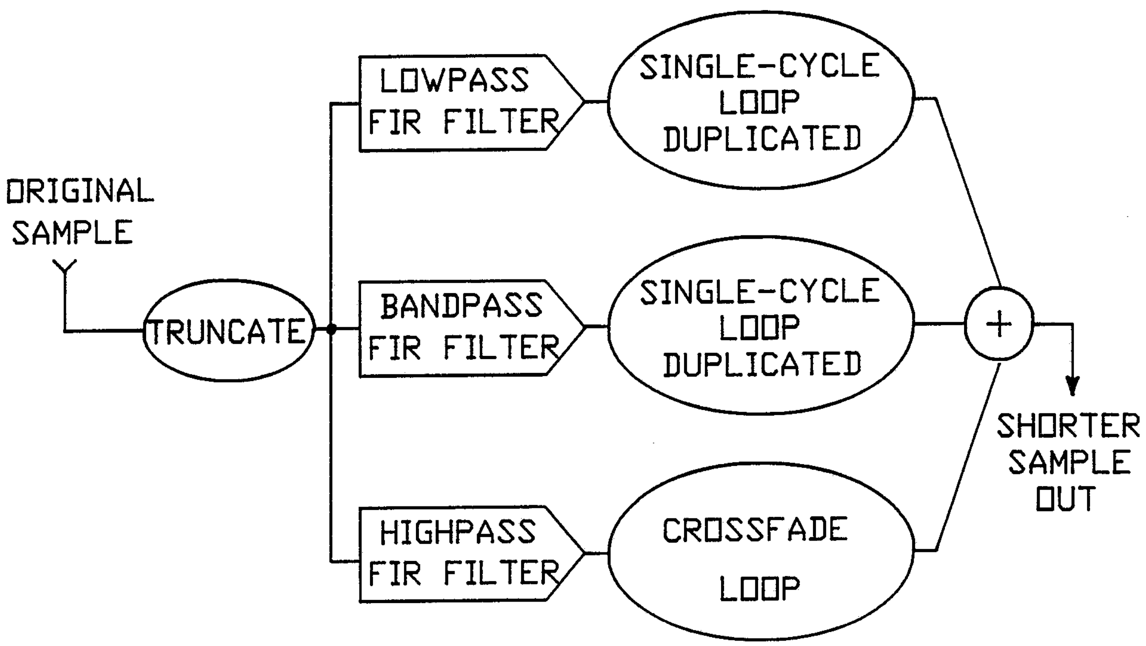

A 440 Hz
dB
0
-10
-20
-30
-40
-50
-60
0 1 2 3 4 KHz
A 440 Hz AS SAMPLED
F 349 Hz
A 440 Hz TRANSPOSED DOWN TO F 349 Hz
F 349 Hz AS SAMPLED
A 440 Hz
F 349 Hz TRANSPOSED UP TO A 440 Hz
COMBINED A 440 Hz AS SAMPLED AND F 349 Hz TRANSPOSED UP TO A 440 Hz
COMBINED A 440 Hz AS SAMPLED AND F 349 Hz TRANSPOSED UP TO A 440 Hz

FIG.-12



DIGITAL FINITE IMPULSE RESPOSE FILTER. THE FILTER COEFFICIENTS, C_I, MUST ALL BE REAL TO INSURE A LINEAR PHASE RESPONSE. THE ORDER OF THE FILTER IS THE NUMBER OF STAGES, N.

FIG.—13



ORIGINAL SAMPLE IS TRUNCATED, BANDSPLIT (IN THIS CASE, INTO 3 BANDS), SEPARATELY LOOPED, AND THEN RECOMBINED. RESULT IS A MUCH SHORTER SAMPLE. ALL BANDSPLIT FILTERS MUST BE OF THE SAME ORDER TO INSURE PHASE CONSISTENCY UPON RECOMBINATION.

FIG.—14

INTERNATIONAL SEARCH REPORT

International Application No PCT/US91/00223

I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all) [6]

According to International Patent Classification (IPC) or to both National Classification and IPC

IPC (5) : G10H 1/02

U.S. CL. 84/603-607, 622-625, 627,639-643, Digests 9 and 26

II. FIELDS SEARCHED

Minimum Documentation Searched [7]

Classification System	Classification Symbols
U.S.	84/603, 625, 627

Documentation Searched other than Minimum Documentation to the Extent that such Documents are Included in the Fields Searched [8]

III. DOCUMENTS CONSIDERED TO BE RELEVANT [9]

Category *	Citation of Document, [11] with indication, where appropriate, of the relevant passages [12]	Relevant to Claim No [13]
X, P	US, A, 4,916,996 (SUZUKI ET AL). 17 APRIL 1990, See the entire document.	1-12
X	US, A, 4,633,749 (FUJIMORI ET AL.) 06 JANUARY 1987, See the entire document.	1-12

* Special categories of cited documents: [10]

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

IV. CERTIFICATION

Date of the Actual Completion of the International Search	Date of Mailing of this International Search Report
14 FEBRUARY 1991	16 APR 1991
International Searching Authority	**Signature of Authorized Officer**
ISA/US	Veronica Lise for STANLEY J. WITKOWSKI

Form PCT/ISA/210 (second sheet) (Rev.11-87)