



(12) **EUROPÄISCHE PATENTANMELDUNG**

(43) Veröffentlichungstag:
09.09.2009 Patentblatt 2009/37

(51) Int Cl.:
G10H 1/06 (2006.01)

(21) Anmeldenummer: **09001575.1**

(22) Anmeldetag: **05.02.2009**

(84) Benannte Vertragsstaaten:
AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO PL PT RO SE SI SK TR
Benannte Erstreckungsstaaten:
AL BA RS

(72) Erfinder: **Neubäcker, Peter**
81371 München (DE)

(74) Vertreter: **Hausfeld, Norbert**
Schaefer Emmel Hausfeld
Patentanwälte
Krohnstieg 43
22415 Hamburg (DE)

(30) Priorität: **07.03.2008 DE 102008013172**

(71) Anmelder: **Neubäcker, Peter**
81371 München (DE)

(54) **Verfahren zur klangobjektorientierten Analyse und zur notenobjektorientierten Bearbeitung polyphoner Klangaufnahmen**

(57) Die Erfindung betrifft ein Verfahren zur klangobjektorientierten Analyse und zur notenobjektorientierten Bearbeitung einer polyphonen, digitalisierten Klangaufnahme, die als Zeitsignal $F(A, t)$ vorliegt, umfassend die folgenden Analyse- und Bearbeitungsschritte: abschnittsweises Auslesen des Zeitsignals $F(A, t)$ unter Verwendung einer Fensterfunktion und sich überlappenden Fenster; Fourier-Transformation des ausgelesenen Signals in den Frequenzraum, insbesondere durch Ausführen einer diskreten Fourier-Transformation; Berechnen eines Energiewertes E zu jedem Bin aus der Frequenzamplitude, die sich aus der Fourier-Transformation ergibt, insbesondere durch Quadrieren der Real- und Imaginärteile oder Bildung eines davon abgeleiteten Energiewertes; Erzeugen einer Funktion $F(t, f, E)$; Identifizierung von Ereignisobjekten; Identifizierung von Noten-

objekten; Vergleich des zeitlichen Auftretens von Ereignisobjekten und Notenobjekten und Zuordnung von Ereignisobjekten zu Notenobjekten für den Fall zeitlich plausiblen Auftretens; Berechnen von spektralen Anteilsfaktoren zu jedem Notenobjekt; Zuordnen von Signalanteilen des Frequenzsignals $F(f, t, E)$ zu gefundenen Notenobjekten anhand der berechneten Anteilsfaktoren; Rücktransformation der einem Notenobjekt zugeordneten Frequenzsignalanteile in ein Zeitsignal; graphische Darstellung der Notenobjekte und/oder Ereignisobjekte in einer Zeit/Frequenzdarstellung auf einem Monitor; anwendergesteuerte oder automatisierte Bearbeitung eines oder mehrerer Notenobjekte; Speicherung der Zeitsignale bearbeiteter Notenobjekte; Wiedergabe der gespeicherten Zeitsignale bearbeiteter Notenobjekte zusammen mit dem um das einem Notenobjekt zugeordnete Zeitsignal verminderte Zeitsignal.

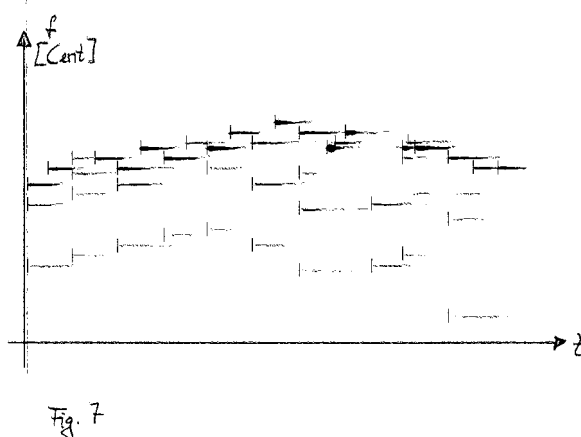


Fig. 7

Beschreibung

[0001] Die vorliegende Erfindung betrifft ein Verfahren zur klangobjektorientierten Analyse und zur notenobjektorientierten Bearbeitung von polyphonen Klangaufnahmen nach Anspruch 1.

[0002] Es ist seit langem bekannt, Tonaufnahmen mit musikalischem Inhalt einer klanglichen Nachbearbeitung zu unterziehen. Fand dies in der Studioteknik früherer Jahre noch unter Zuhilfenahme von kostspieligen Hardwarekomponenten statt, wie z.B. mit komplexen Filterbänken, so werden dazu heutzutage Computer und spezielle Computerprogramme eingesetzt, die weit weniger kostenintensiv sind und daher eine größere Verbreitung gefunden haben. Ein Übriges zu dieser Entwicklung hat der Einzug digitaler Aufnahmetechnik getan. Ziel solcher Nachbearbeitungen ist es in der Regel, den Klang der Aufnahmen zu verbessern oder Klangeffekte einzuarbeiten. Dieser Art von klanglicher Nachbearbeitung ist gemein, dass sie rein effektorientiert arbeitet und den musikalischen Gehalt des Signals nicht zu erkennen vermag, sondern das Audio-Signal nur als eine sich in der Zeit verändernde Signalamplitude versteht.

[0003] Im Stand der Technik sind z.B. aus der EP 0 750 776 B1 bzw. der DE 696 14 938 T2 ein Verfahren und eine Vorrichtung zur Änderung des Klanges und der Tonhöhe von Audio-Signalen bekannt. Es wird als nachteilig angesehen, dass dort nicht mit komplexem Klangmaterial gearbeitet werden kann, wie es bei einer üblichen Musikproduktion aber vorliegt.

[0004] Wünschenswert ist eine Bearbeitung von Audio-Material auf der Ebene der Einzelnoten, aus dem die Klangaufnahme besteht. Es ist im Stand der Technik bekannt, aus einer Audio-Aufnahme Einzelnoten im Hinblick auf deren Notenhöhe, Notenlänge und Auftrittszeitpunkt zu extrahieren. Eine solche Notenextraktion ist z.B. aus der DE 10 2004 049 477 A1 bekannt, um eine Melodielinie aus einem Audio-Signal zu bestimmen. Aus der WO 02/084641 A1 ist die Überführung eines Audio-Signals in eine notenbasierte Beschreibung bekannt, um das Audio-Signal in einer Datenbank referenzieren zu können. Eine Bearbeitung der extrahierten Noten, z.B. durch Änderung in der Frequenz oder Verschieben in der Zeit, findet nicht statt. In diesen Schriften wird weiterer Stand der Technik genannt.

[0005] Ein besonders kritischer Punkt bei Bearbeitungen von Audio-Material ist, dass der ursprüngliche Klangeindruck z.B. einer Gesangsstimme auch nach der Bearbeitung erhalten bleiben soll. Dies gelingt in hervorragender Weise in der zum Stand der Technik gehörenden Software "Melodyne" der Celemony Software GmbH, die auf einer notenbasierten Herangehensweise fußt. Allerdings setzt diese Software voraus, dass einstimmiges Material vorliegt. Akkord-Instrumente wie Gitarre, Klavier oder Chorgesang lassen sich bisher nicht zufriedenstellend tonbasiert bearbeiten. Bisher konnten solche Akkordiaufnahmen nur akkordweise geschnitten oder per Timestretching in der Zeit oder Tonhöhe bear-

beitet werden, allerdings ohne Zugriff auf die einzelnen Töne eines Akkords. So war es bisher nicht möglich, einen einzelnen Akkordton (z.B. das E eines C-Dur-Akkords) zu verändern (z.B. auf Es für C-Moll), ohne gleichzeitig die anderen Töne des Akkords mit zu bearbeiten.

[0006] Akkord-Erkennung und Ansätze zur Identifikation von Einzelnoten sind im Stand der Technik allerdings bereits bekannt, z.B. um Noten drucken zu können (Software WIDI) oder automatisch Titel erkennen zu können (DE 10 2004 049 477 A1).

[0007] Es ist die Aufgabe der vorliegenden Erfindung einen Weg aufzuzeigen, wie eine notenobjektorientierte Bearbeitung polyphonen Klangmaterials erfolgen kann. Diese Aufgabe wird mit einem Verfahren nach Anspruch 1 gelöst. Vorteilhafte Ausgestaltungen sind in den Unteransprüchen angegeben.

[0008] Das erfindungsgemäße Verfahren identifiziert in einem ersten Schritt automatisch in einem aufgenommenen Audio-Material musikalische Objekte im Sinne von Noten. Diese Objekte werden dann in einem zweiten Schritt aus der Gesamtaufnahme klanglich extrahiert und dadurch ohne merklichen Klangverlust oder merkliche Klangverfälschung manipulierbar gemacht. Diese klangliche Extraktion ist im Stand der Technik nicht zu finden und bedeutet, daß einem identifizierten Notenobjekt ein Teil des Gesamtklanges zugeordnet wird. Dadurch kann die ursprüngliche Aufnahme sogar musikalisch in der Weise frei verändert werden, dass die Noten ihre relative Position zueinander in Tonhöhe und Zeit ändern können, der ursprüngliche Klangeindruck bleibt aber erhalten. Das Verfahren sieht weiter vor, daß die identifizierten Notenobjekte dem Anwender zur Bearbeitung zur Verfügung gestellt werden. Der Anwender kann dann einzelne oder mehrere musikalische Objekte verändern, z.B. in der Tonhöhe verschieben. Nach einer solchen Bearbeitung durch einen Anwender erfolgt die klangliche Wiedergabe, die Resynthese, indem das veränderte Objekt zusammen mit den nicht veränderten Objekten bzw. mit dem um das veränderte Objekt verminderten Gesamtsignal wiedergegeben wird. Das Ausgangsmaterial kann dabei bereits in digitaler Form oder auch als analoges Signal vorliegen. Vor der Analyse eines analogen Signals muß dann ggf. eine Digitalisierung durchgeführt werden.

[0009] Das Verfahren kann verschiedene Anwendungen finden. Es lassen sich z.B. gezielt einzelne Verspieler beseitigen: ein Pianist hat versehentlich einen Ton zuviel angeschlagen, der in der Nachbearbeitung entfernt wird. Eine andere Anwendung betrifft das Nachstimmen, also die Korrektur für eine verstimmte Gitarre oder einen unsauberen Streichersatz. Temperierte Aufnahmen können auf reine Stimmung übertragen werden. Es lassen sich Aufnahmen Umharmonisieren, etwa das Riff einer Gitarre von C-Dur in F-Moll. Bisher konnte ein Akkord nur insgesamt in der Tonhöhe verschoben, nicht aber die harmonische Beziehung seiner Einzeltöne verändert werden. Die Möglichkeiten reichen sogar bis hin zu einer Neukomposition, denn es besteht Zugriff auf die

Einzelnote.

[0010] Das erfindungsgemäße Verfahren richtet sich vorwiegend an aufgenommene Einzelspuren während der Musikproduktion. Es ist aber auch anwendbar auf fertig gemischte Titel, um diese in einem neuen musikalischen Gewand erscheinen zu lassen.

[0011] Mit der bisherigen Technik war das oben Gesagte nur für klanglich monophones Material realisierbar, d.h. zur Bearbeitung von beispielsweise isoliert aufgenommenen Gesangs- oder Bläserstimmen, bei denen nur je eine Note mit identifizierbarer Tonhöhe gleichzeitig erklingt. Mit dem erfindungsgemäßen Verfahren wird das Ziel der Einzelnotenbearbeitung realisierbar für polyphon gespielte Instrumente, also solche, bei denen in der Regel mehrere Noten oder ganze Akkorde gleichzeitig erklingen, wie Klavier, Gitarre usw., wobei nicht nur der Akkord als Ganzes transponierbar wird (d.h. Veränderung der Tonhöhe unter Beibehaltung der relativen Tonhöhenbezüge innerhalb eines Akkordes), sondern auch und besonders die Noten innerhalb eines Akkordklanges relativ zueinander verändert werden können, wie etwa die Veränderung eines Akkordklanges von Dur zu Moll.

[0012] Das Verfahren bezieht sich vordringlich auf bereits aufgenommenes musikalisches Material, also nicht auf die Analyse und Bearbeitung in "Echtzeit", d.h. im Moment der musikalischen Aufführung. Das hat damit zu tun, dass das erfindungsgemäße Verfahren eine sinnvolle Erkennung von "Noten" im Sinne von abgeschlossenen Objekten in einem gewissen Kontext voraussetzt. Daher kann die Analyse zwar bei einer laufenden Aufnahme im Hintergrund durchgeführt werden, es muss dabei aber auf einen bereits aufgenommenen Zeitzusammenhang in der Größenordnung z.B. von einigen Sekunden zugreifen können.

[0013] Das Verfahren ist dazu ausgelegt, einzelne Tonhöhen oder Tonhöhenverläufe im Gesamtsignal zu finden und zu unterscheiden, nicht aber einzelne Klangquellen zu trennen. Es geht also nicht darum, z.B. aus einer Aufnahme der Geräusche an einer Straße oder mehrerer Sprecher in einem Raum die einzelnen Quellen der Geräusche oder Stimmen zu trennen. Es werden in der Regel zwei Noten, die auf gleicher Tonhöhe von zwei verschiedenen Instrumenten gleichzeitig gespielt wurden, als ein einziges Objekt identifiziert, ebenso der Klang von vielen Ersten Geigen im Orchester, die dieselbe Note spielen, als eine einzige Note. Der hier überwiegend verwendete Begriff des Notenobjektes unterstreicht, dass mit "Noten" im Sinne dieser Erfindung nicht Noten im eigentlichen musikalischen Sinne gemeint sein müssen, wenngleich ein identifiziertes Notenobjekt einer Note im eigentlichen musikalischen Sinn entsprechen kann, aber nicht muß.

[0014] Zu findende Noten müssen nicht, wie z.B. eine Klavierton, eine über die Zeit konstante Tonhöhe haben, sondern können auch, wie z.B. eine Singstimme mit Vibrato und/oder Portamento, über die Zeit einen beliebigen Tonhöhenverlauf haben. Wenn solche Tonhöhenverläufe in sich konsistent sind bleiben die Noten bei Zeit-

verfolgung des Signals erkennbar. Insofern können auch zwei Töne, die sich bei unterschiedlichem Verlauf in der Tonhöhe kreuzen, als zwei verschiedene Noten identifiziert werden.

[0015] Das Verfahren besteht im wesentlichen aus den beiden Hauptschritten a) Identifizierung der am Gesamtklang beteiligten Einzelobjekte, also der Noten und ggf. auch der mehr ereignishaft/perkussiven Klangereignisse, und b) klangliche Zerlegung des Gesamtklanges in die gefundenen Einzelobjekte, deren Summe den Gesamtklang ergibt, deren jedes dadurch aber separat manipulierbar wird, ohne den Klang der übrigen Objekte bzw. des Gesamtklanges im Sinn von unerwünschten hörbaren Artefakten zu beeinflussen. Gerade im Schritt b) unterscheidet sich die vorliegende Erfindung vom Stand der Technik.

[0016] Das Ergebnis des Identifizierungsschrittes a) kann auch für sich genommen ohne den Klangzerlegungsschritt b) verwendet werden, wenn es nur um die Erkenntnis und Darstellung des musikalischen Inhaltes geht, ohne dass in den Klang selbst eingegriffen werden soll. Das kann der Fall sein, wenn man von einer bestehenden Aufnahme eine musikalische Notation erzeugen will oder den musikalischen Inhalt auf andere Weise graphisch veranschaulichen will. Oder es kann dazu dienen, die Folge der musikalischen Harmonien zu erkennen und zu benennen, um z.B. die Musik mit weiteren Instrumenten anzureichern.

[0017] Die Qualität des Ergebnisses von Identifizierungsschritt a) beeinflusst auch die Qualität der klanglichen Zerlegung in Schritt b). Wenn Schritt a) nur wie erwähnt zur Ermittlung des musikalischen Inhaltes dienen soll, genügt es in der Regel zu ermitteln, dass zu einer bestimmten Zeit ein Klangobjekt auftritt, das z.B. die Tonhöhe der Note "Cis" hat, und wie lange dieses Objekt klingt. Wenn hingegen die klangliche Zerlegung in Schritt b) das Ziel ist, dann ist es vorteilhaft so viele Informationen wie möglich über den Verlauf und die Parameter der gefundenen Notenobjekte auszuwerten, wie z.B. den genauen Verlauf der Tonhöhenkurve in der Zeit, die Amplitude des Objektes und ihre Änderung in der Zeit, die Art des Einsetzens einer Note, die Konsistenz des Gemisches ihrer Teiltöne im Verhältnis zu Noten mit ähnlicher Tonhöhe in derselben Aufnahme, usw.. Je nach Anspruch kann man aber auch auf bestimmte Informationen verzichten.

[0018] Wie beschrieben ist die Haupteigenschaft der im Audio-Material zu findenden Einzelobjekte, dass sie eine konsistente Tonhöhe bzw. einen konsistenten Tonhöhenverlauf haben. Des Weiteren sollten sie einen konsistenten Verlauf ihrer einzelnen Teiltöne aufweisen. Das heißt, dass angenommen wird, dass ein zu findendes Klangobjekt aus dem Verlauf eines Grundtones besteht, sowie aus einer im Prinzip unbegrenzten Zahl von Obertönen, die annähernd ganzzahlige Vielfache der Grundtonfrequenz sein sollen. Weiter sollte der Verlauf der zu einem Klangobjekt gehörenden Teiltöne keine unmotivierten Sprünge aufweisen. Diese Annahmen sind aus

den Eigenschaften der von natürlichen Instrumenten erzeugten Klänge abgeleitet. Daher hat das erfindungsgemäße Verfahren möglicherweise dort seine Grenzen, wo Musik auf beliebige willkürliche Weise synthetisch erzeugt wurde.

[0019] Die im Audio-Material zu identifizierenden und oben beschriebenen Objekte kann man als "klanghafte" Objekte bezeichnen, in den Ansprüchen werden sie als Notenobjekte benannt. Deren Haupteigenschaft es ist, eine Tonhöhe oder einen Tonhöhenverlauf über eine wahrnehmbare Dauer zu haben, und die Kurvenform ihres Zeitsignals verhält sich im Wesentlichen periodisch oder quasiperiodisch. Es werden davon die nicht klanghaften Objekte unterschieden, die geräuschhaften Objekte. Ereignishafte Objekte sind eine Untermenge der geräuschhaften Objekte.

[0020] Die Ereignisobjekte zeichnen sich dadurch aus, dass sie im Zeitsignal einen sprunghaften Amplitudenanstieg aufweisen und schon deshalb an dieser Stelle nicht periodisch sind. Außerdem klingen Sie meist schnell aus. Diese Objekte werden musikalisch auch meist von einem plötzlichen Ereignis erzeugt, wie vom Anschlagen oder Anzupfen einer Saite, oder dem Anschlagen eines Schlaginstrumentes wie einer Trommel. In diesem Zusammenhang können sie vom erfindungsgemäßen Verfahren auch noch weiter unterschieden werden: falls auf ein ereignishaftes Objekt sofort der Amplitudenanstieg eines klanghaften Objektes folgt, kann angenommen werden, dass das Ereignis den Anschlag einer Note darstellt und kann deshalb dieser Note zugeordnet werden. In diesem Fall können das klanghafte und das ereignishafte Objekt zur Darstellung und Handhabung zu einem einzigen Objekt zusammengefasst werden, das ebenfalls als Notenobjekt zu bezeichnen ist. Zur Klangerzeugung bei der Resynthese, d.h. bei der Klangwiedergabe nach der Bearbeitung eines Notenobjektes oder mehrerer Notenobjekte, kann es aber sinnvoll sein, die beiden genannten Teilobjekte aus klanglichen Gründen getrennt zu verarbeiten.

[0021] Wenn das ereignishafte Objekt nicht einem folgenden klanghaften Objekt zugeordnet werden kann, kann angenommen werden, dass es sich um eine rein perkussiv-rhythmisch verstandene Note ohne Tonhöhenbedeutung handelt, wie etwa ein Trommelschlag. Dieses kann in der weiteren Bearbeitung dementsprechend anders behandelt werden.

[0022] Die geräuschhaften Objekte haben mit den ereignishaften Objekten gemeinsam, dass sie keinen periodischen Klanganteil haben, unterscheiden sich aber von diesen dadurch, dass sie nicht plötzlich anfangen und auch nicht schnell ausklingen, sondern zeitlich ausgedehnter sein können. Solche Objekte können z.B. Zischlaute der Konsonanten im Gesang, Atemgeräusche, Nebengeräusche des Spielers wie ein Fingerrutschen auf der Gitarrensaite, oder auch gar nicht zur eigentlichen Musik gehörende Nebengeräusche sein. Solche geräuschhaften Objekte könnten z.B. dem Anwender separat zur weiteren Manipulation zur Verfügung ge-

stellt werden oder einfach pauschal als "Restsignal" behandelt werden.

[0023] Nachfolgend soll das erfindungsgemäße Verfahren an einem Ausführungsbeispiel erläutert werden unter Bezugnahme auf die Figuren, in denen Ergebnisse einzelner Verfahrensschritte dargestellt sind. Das Verfahren wurde durchgeführt an einem 6-Sekunden-Ausschnitt aus einer Klavieraufnahme der Invention No. 1, C-Dur von J.S. Bach. Es zeigen:

Fig. 1 das Audio-Signal $F(A,t)$ als Amplitude A über die Zeit t ;

Fig. 2 einen vergrößerten Zeitausschnitt aus der Fig. 1 über 0.5 Sekunden;

Fig. 3 nach Transformation des Audio-Signals der Fig. 1 in den Frequenzbereich die grautonkodierte Energien der einzelnen Bins in ihrer Augenblicksfrequenz, wobei dunkle Bereiche hohe Energie bedeuten, in einer Darstellung der Frequenz f in Cent über der Zeit t : $F(f, t, E)$;

Fig. 4 eine Darstellung gemäß Fig. 3 mit einem Schnitt in x - und y -Richtung, der die errechneten Energien E bei dieser Frequenz f_1 und der Zeit t_1 zeigt;

Fig. 5 eine Energielandschaft $F(f, t, E)$ ähnlich der in Fig. 4 gezeigten, bei der die Energiewerte über eine Fensterfunktion aufsummiert und in Zeitrichtung geglättet wurden, mit einem Schnitt in x - und y -Richtung bei der Frequenz f_1 und der Zeit t_1 ;

Fig. 6 eine Relevanzlandschaft zu der Energielandschaft der Fig. 5 mit einem Schnitt in x - und y -Richtung mit dem ersten gefundenen Maximum; und

Fig. 7 alle gefundenen Notenobjekte nach Abschluss einer iterativ ausgeführten Identifizierung.

[0024] Es werden zur Erläuterung der Erfindung Begriffe verwendet die zunächst in ihrer Bedeutung definiert werden sollen. Die Begriffe "Zeitbereich" bzw. "Frequenzbereich" entsprechen dem üblichen Gebrauch von time domain und frequency domain im Englischen, und betreffen also Untersuchungen oder Berechnungen entweder im ursprünglichen Zeitsignal $F(A, t)$ selbst (= Zeitbereich) oder in seiner Repräsentation $F(f, t)$ in der über eine diskrete Fourier-Transformation, speziell FFT, überführten Form (= Frequenzbereich).

[0025] "Fensterfunktionen" dienen zum Ein- und Ausblenden eines Signals bei seiner Bearbeitung an einem Zeit- oder Frequenzort. Sie können sich also auf die Zeit oder auf das Frequenzspektrum beziehen. Die Fensterform ist nicht festgelegt und kann in einem einfachen Fall

ein Dreiecksfenster sein. Bessere Ergebnisse ergeben z.B. ein von-Hann-Fenster. Die Wahl der Fensterform kann für den jeweiligen Zweck optimiert werden. Für das nachfolgend beschriebene Ausführungsbeispiel eines erfindungsgemäßen Verfahrens werden sich überlappende Fenster verwendet.

[0026] Als "Bins" werden die Frequenzbänder bezeichnet, die sich aus der FFT ergeben. Damit werden auch die möglicherweise geänderten Frequenzen eines Bandes bezeichnet, wenn die Methode der Augenblicksfrequenz angewendet wird.

[0027] Die Augenblicksfrequenz eines Bins ergibt sich aus der Berücksichtigung des Phasenwertes des jeweiligen Bins. Wenn die Analysefenster sich überlappen, kann aus der Differenz zwischen dem durch den Zeitfortschritt erwarteten und dem tatsächlich vorhandenen Phasenwert des Bins die Augenblicksfrequenz des jeweiligen Bins ermittelt werden. Je häufiger die Überlappung ist, desto mehr benachbarte Bins können eine bestimmte Frequenz repräsentieren, die nicht mit der rechnerischen Frequenz des Bins selbst übereinstimmen muss.

[0028] Es werden für das erfindungsgemäße Verfahren für die Bins energiehafte Größen berechnet, die die Kurznotation E erhalten und nachfolgend als "Energie" und "Magnitude" bezeichnet werden sollen. Nach der FFT werden jedem Bin jeweils eine Energie zugeordnet, die sich aus Real- und Imaginärteil der Fourier-Reihe nach $(\text{Re} \cdot \text{Re}) + (\text{Im} \cdot \text{Im})$ errechnet, so dass ein mit der Amplitude der Frequenz zusammenhängender Wert entsteht. Die Magnitude ist die Wurzel daraus. Zur optimalen Skalierung des jeweiligen Wertes in der Auswertung kann aber der Magnitudenwert ggf. mit einem geeigneten Wert zwischen 1 und 2 potenziert werden, so dass die dann erhaltene Größe etwas zwischen Energie und Magnitude darstellt. Der Begriff Energie wird also hier in einem allgemeineren Sinn verwendet, und nicht im Sinne von Schallenergie oder anderer Energiebegriffe.

[0029] Das Cent ist ein im musikalischen Zusammenhang verwendetes Maß für Frequenzverhältnisse, also Intervallgrößen, definiert als $\text{Cent} = \log(f_1/f_2) / \log(2) \cdot 1200$. Deshalb ist in dieser Einheit ein musikalisches Intervall unabhängig von seiner absoluten Tonhöhe immer gleich groß, nämlich Halbton = 100 Cent, Oktave = 1200 Cent.

[0030] Zur Identifizierung von Notenobjekten bedient sich das beschriebene Verfahren bildhaft gesprochen einer Energielandschaft, worunter ein dreidimensionales mathematisches Gebilde $F(t, f, E)$ verstanden wird, bei dem die x-Achse die Zeit t , die y-Achse die Frequenz f und die z-Achse eine Energie E am jeweiligen Zeit/Frequenz-Ort darstellen. Dabei ist zweckmäßig die Frequenzachse nach Cent skaliert, damit Tonintervalle in jedem Frequenzbereich immer gleich groß sind. Die Energielandschaft ist im weiter zu beschreibenden Verfahren durch diskrete Punkte repräsentiert, die Zeitachse durch die Messpunkte in der Zeit in z.B. ca. 0.01 s Abstand, die Frequenzachse durch Punkte im Intervallabstand von z.B. je 5 Cent. In einigen Verfahrensschritten

werden die diskreten Punkte über Fensterfunktionen in kontinuierliche Spektrogramme gewandelt, was aber als optional zu betrachten ist.

[0031] Das untersuchte Audiomaterial liegt nach Aufnahme und Analog-DigitalWandlung z.B. als Audiodatei im PCM-Format vor (Pulse-Code-Modulation, werte- und zeitdiskretes Signal). Größenangaben im weiteren Text, wie z.B. für Analysefenster, beziehen sich auf ein digitales Signal, das mit einer Samplerate von 44100 Samples/s vorliegt. Für andere Sampleraten wären die Größen entsprechend anzupassen.

a) Identifizierung von Notenobjekten und Ereignisobjekten

[0032] Das exemplarisch beschriebene Verfahren arbeitet sowohl bei der Analyse als auch bei der Klangextraktion für bestimmte Teilaufgaben direkt im Zeitbereich (Fig. 1 und 2), für andere im Frequenzbereich (Fig. 3 bis 7). Dabei eignet sich die Behandlung im Zeitbereich besser für die ereignishaften Objekte, die Behandlung im Frequenzbereich besser für die klanghaften Objekte. Da nach diskreten Notenobjekten in der Zeit gesucht wird, wird das Signal nicht kontinuierlich verarbeitet, sondern es wird sowohl im Zeitbereich als auch im Frequenzbereich immer ein Zeitabschnitt zwischengespeichert, der dann untersucht wird.

[0033] Fig. 1 zeigt einen Signalverlauf $F(A, t)$ zu einer Audio-Aufnahme. Figur 2 zeigt einen vergrößerten Ausschnitt dazu. Mit 1 ist in diesen Figuren ein Ereignisobjekt bezeichnet, daß sich an einem Amplitudensprung festmachen läßt. Der mit 2 bezeichnete Bereich zeigt demgegenüber einen deutlich periodischeren Verlauf.

[0034] Für die Bearbeitung im Frequenzbereich wird das Signal mit gleichmäßig aufeinander folgenden und sich überlappenden Fensterfunktionen ausgelesen und zunächst über eine FFT in ein komplexes Array für die jeweilige Zeitscheibe überführt. Die Größe der FFT kann z.B. 2048 Samples betragen, die Überlappung sollte mindestens 4-fach sein. Es ergeben sich so Abstände der Zeitscheiben von z.B. 512 Samples oder rund 0.01 sec.

[0035] Ein Beispiel für ein weiter bearbeitetes Transformationsergebnis zeigt Figur 3. Es ist dort eine Funktion $F(f, t, E)$ dargestellt, die sich aus der Fourier-Transformation des in den Figuren 1 und 2 gezeigten Zeitsignals $F(A, t)$ dadurch ergibt, dass zu den Bins die Augenblicksfrequenzen und Energien ermittelt werden. Als Grautöne kodiert sind die Energien der einzelnen Bins in ihrer Augenblicksfrequenz über der Zeit aufgetragen. Ereignisobjekte 3 zeichnen sich in dieser Darstellung durch die gleichmäßige Aufteilung der Augenblicksfrequenzen aus, die die Geräuschhaftigkeit dieses Objektes anzeigt. Notenobjekte zeichnen sich durch die Konzentration der Energie auf wenige Augenblicksfrequenzen aus. Dieser Unterschied ist dadurch bedingt, dass ein im Grunde nicht-periodischer Signalverlauf nur durch eine Überlagerung einer großen Zahl von periodischen Funktionen adäquat dargestellt werden kann, während sich

periodische Bereiche gut durch einige wenige periodische Funktionen abbilden lassen. Periodische Bereiche im Zeitsignal führen deshalb zu einer energiestarken Überlagerung benachbarter Bins 4 auf einer gemeinsamen Augenblicksfrequenz.

[0036] Zu dem Signal $F(f, t, E)$ im Frequenzbereich werden die folgenden Werte bestimmt: die Magnituden aller Bins, die Augenblicksfrequenzen aller Bins, die Tonalitätswerte aller Bins. Der Tonalitätswert ist dabei eine rechnerische Größe, die für den Grad der Periodizität in der Binfrequenz steht. Er wird für je ein Bin berechnet, indem ermittelt wird, wie nah die Augenblicksfrequenzen der Nachbarbins an der Augenblicksfrequenz des in Frage stehenden Bins liegen. Die Anzahl der einbezogenen Nachbarbins ist dabei gleich der Anzahl der Fensterüberlappungen, da diese bestimmt, wie viele Bins eine Frequenz repräsentieren können. Der Tonalitätswert eines Bins wird umso höher, je näher die Augenblicksfrequenzen der Bins in seiner Umgebung zusammen liegen. Dabei bedeutet ein hoher Tonalitätswert, dass tendenziell ein Notenobjekt vorliegt, während ein niedriger Tonalitätswert tendenziell für ein Ereignisobjekt spricht. Die Tonalitätswerte werden auf einen Wertebereich zwischen 0 und 1 skaliert. Zusätzlich wird jedem Bin ein Geräuschhaftigkeitswert zugewiesen, der direkt aus dem Tonalitätswert abgeleitet wird, und sich berechnet als $1 - \text{Tonalitätswert}$. Ein plötzlicher Anstieg der Geräuschhaftigkeitswerte spricht für ein Ereignisobjekt.

[0037] Es wird dann eine Energielandschaft erzeugt, die exemplarisch in Fig. 4 als Schnitt entlang der Zeit t_1 und der Frequenz f_1 dargestellt ist, und die die Verteilung der Energie E in der Zeit t und Frequenz f repräsentiert und dem eigentlichen Finden der Notenobjekte als Höhenzüge in dieser Landschaft dient. Eine Glättung in Zeitrichtung kann durchgeführt werden. Weiterhin wird für die weiteren Berechnungen bevorzugt eine durch Einblenden der Energiewerte über eine Fensterfunktion aus der Energielandschaft der Figur 4 gewonnene modifizierte Energielandschaft verwendet. Es ergibt sich dann eine in Fig. 5 dargestellte geglättete Funktion $F(f, t, E)$. In beiden Figuren treten Objekte 5 deutlich hervor, denen hohe Energiewerte zukommen.

[0038] Da es zunächst um das Auffinden klanghafter Objekte geht, werden die errechneten Energien der Bins zusätzlich mit den ermittelten Tonalitätswerten gewichtet: für jede Zeitscheibe wird für jedes Bin seine Energie mit seinem Tonalitätswert multipliziert. Diese vorgenommene Wichtung verändert das Ergebnis nur graduell, kann deshalb auch weggelassen werden. Gemäß der Augenblicksfrequenz des Bins wird dann seine Cent-Position (= y-Position) in der Landschaft ermittelt und von diesem Punkt ausgehend das Energie-Tonalitätsprodukt mit einer gewissen Verteilungsbreite in Cent-Richtung über eine Fensterfunktion auf die Landschaft aufsummiert. Die Breite des Verteilungsfensters in Cent liegt zweckmäßig etwa in der Größenordnung von einem Halbton. Der Darstellung der Fig. 4 lag eine solche Gewichtung mit den Tonalitätswerten bereits zu Grunde.

Wenn alle Zeitscheiben in der Landschaft aufsummiert sind, kann die Landschaft in Zeitrichtung mit einem Tiefpassfilter geglättet werden (siehe Figur 5). Dadurch wird das Auffinden von zusammenhängenden Notenobjekten als Höhenzüge erleichtert. Die Gesamtenergie der Landschaft wird aufsummiert und steht als Wert für das Abbruchkriterium der nachfolgenden Iteration zur Verfügung.

[0039] Das Identifizieren der Notenobjekte geschieht durch eine Iteration in der Weise, dass das jeweils am deutlichsten hervortretende Objekt im Sinne des Maximums als Höhenzug verfolgt wird und dessen Energie anschließend aus der Landschaft subtrahiert wird, worauf das nächste am deutlichsten hervortretende Objekt gesucht wird, usw. Das klanglich am deutlichsten hervortretende Objekt ist aber nicht identisch mit dem höchsten Höhenzug in der Energielandschaft $F(f, t, E)$. Das ist dadurch bedingt, dass ein klanghaftes Objekt nicht definiert ist durch einen einzelnen Höhenzug in der Energielandschaft, sondern dass angenommen werden muss, dass die Energie in Höhenzügen auf den ganzzahligen Vielfachen einer Grundfrequenz auch zu eben diesem gesuchten klanghaften Objekt gehören, weil sie Obertöne zu dem Grundton mit der Grundfrequenz sind. Dabei kann es durchaus sein, dass die Energie des Grundtons schwächer ist als die der höheren Teiltöne, trotzdem aber das Objekt auf der Frequenz des Grundtons gefunden und verfolgt werden soll. Hintergrund dieser Überlegungen ist, daß die Grundfrequenz zwar die Tonhöhe eines Tones bestimmt, der Klang eines Tones wird aber maßgeblich von den Obertönen geprägt.

[0040] Um dem Rechnung zu tragen wird bildhaft gesprochen eine zweite Landschaft erzeugt, die sogenannte Relevanzlandschaft, die exemplarisch in Fig. 6 dargestellt ist, und die in ihren x- und y-Achsen und ihrer Größe mit der Energielandschaft $F(f, t, E)$ identisch ist, ihre z-Werte E' aber aus dieser ableitet: dazu wird für jeden x-y-Koordinatenpunkt der Relevanzlandschaft $F'(f, t, E')$ als z-Wert E' die Summe aller z-Werte E gebildet, die sich in der Energielandschaft $F(f, t, E)$ an diesem x-y-Punkt befinden und an allen Punkten, die den ganzzahligen Frequenzvielfachen des Ausgangspunktes entsprechen. Dabei ist es zweckmäßig, mit zunehmender Ordnungszahl der Frequenzvielfachen den dort vorhandenen Energiewert mit abnehmendem Gewicht zu addieren. Auf diese Weise ergibt sich eine Relevanzlandschaft $F'(f, t, E')$, in der die möglichen Grundtöne mit ihren Teiltönen berücksichtigt sind, und deren höchster Punkt der klanglich relevanteste Punkt der relevantesten Note ist. Die in Figur 6 gezeigte Relevanzlandschaft $F'(f, t, E')$ zeigt die gleichen energetisch hervortretenden Objekte 5. Es hat sich durch die Berücksichtigung der in den Obertönen enthaltenen Energien gegenüber der Figur 5 eine Verschiebung in den relativen Energiehöhen ergeben.

[0041] Das Auffinden der Notenobjekte in der wie eben geschildert erzeugten Relevanzlandschaft, die im Grunde nur eine besondere Energielandschaft ist, nämlich ei-

ne die Obertonenergie berücksichtigende, erfolgt durch ein iteratives mathematisches Verfahren. Es wird nachfolgend erläutert, wie das Auffinden der Notenobjekte in dieser Relevanzlandschaft erfolgt, ohne dass die Erfindung sich darauf beschränkt. Denn das Auffinden der Notenobjekte könnte prinzipiell auch in einer der anderen oben geschilderten oder weiter modifizierten Energielandschaften erfolgen, was aber den Nachteil hätte, dass Obertöne als eigene Noten identifiziert würden und z.B. durch Nachbearbeitung mit den Grundtönen zu verbinden wären. Die Aufgabe, Notenobjekte auch klanglich zu separieren, ist dann gut gelöst, wenn die Verknüpfung zwischen Grund- und Obertönen gelingt. Daher ist die Maximumsuche in der Relevanzlandschaft bevorzugt, weil sie zu den besten Ergebnissen führt.

[0042] Es wird zunächst der höchste Punkt der Relevanzlandschaft gesucht. In Fig. 6 wurde bei t_1 und f_1 das Energiemaximum gefunden. Der Grat des zu diesem Maximum gehörenden Höhenzuges wird in Zeitrichtung vorwärts und rückwärts verfolgt. Dazu wird in der jeweils benachbarten Zeitscheibe der Landschaft das dem zuletzt gefundenen Punkt in Tonhöhenrichtung am nächsten gelegene Maximum in der Landschaft gesucht. Wenn der Abstand zum nächstgelegenen Maximum so groß ist, dass eine Fortsetzung der Tonhöhenlinie als dasselbe Objekt nicht plausibel ist, z.B. bei einem Sprung von mehr als 50 Cent von einer Zeitscheibe zur folgenden, wird die Suche in der aktuellen Richtung abgebrochen. Ebenso wird die Suche abgebrochen, wenn ein bestimmter Höhenwert des gefundenen Maximums unterschritten wird, wie z.B. 10% des Anfangswertes. Die Verfolgung des Höhenzuges geschieht zweckmäßig in der Relevanzlandschaft, weil dessen Verlauf infolge der Mitgewichtung der Teiltöne dem Tonhöhenverlauf des gesuchten Objektes besser entspricht. Wenn die Suche in beide Richtungen abgebrochen ist, wird ein neues Notenobjekt erzeugt und diesem alle Punkte des gefundenen Grates als seinem Tonhöhenverlauf entsprechend hinzugefügt. In den Ansprüchen ist in diesem Zusammenhang davon die Rede, daß ein zum Maximum gehörendes Wertefeld ermittelt wird. Die Art der Ermittlung dieses Wertefeldes kann dabei auch anders als vorab beschrieben erfolgen, indem z.B. andere mathematische Verfahren angewendet werden. Es könnte z.B. in alle Richtungen vom Maximum weg Punkt für Punkt das Wertefeld abgefragt werden, bis in jeder Richtung ein Schwellenwert unterschritten wird. Alle oberhalb der Schwelle liegenden Punkte wären dem Maximum als Wertefeld zuzuweisen.

[0043] Anschließend wird der Energielandschaft $E(f, t, E)$ die rechnerische Energie des gefundenen Notenobjektes entzogen, und zwar an den Orten des Verlaufes seines Grundtones sowie dem aller Teiltöne, also der ganzzahligen Vielfachen der Grundtonfrequenz. Prinzipiell könnte das auch in der Relevanzlandschaft $E'(f, t, E')$ erfolgen, was aber eine schlechtere Ausführungsvariante wäre, weil die mögliche Überlagerung von Obertönen, die zu verschiedenen Grundtönen gehören, bei

Energieentzug aus der Energielandschaft besser aufgehoben wird.

[0044] Mit Vorteil wird aber nicht die gesamte dort vorhandene Energie entzogen, sondern nur ein festgelegter Anteil, z.B. 50%. Dieser Anteil kann z.B. als Parameter vom Anwender auf andere Werte gesetzt werden, weil je nach Audio-Material andere Anteile bessere Ergebnisse ergeben können. Bei starker Überlagerung in den Obertönen kann z.B. eine Herabsetzung auf 25% zu besseren Ergebnissen führen. Der Entzug nur eines Anteils der Energie ist sinnvoll, da zunächst nicht bekannt ist, ob nicht andere gleichzeitig klingende Notenobjekte Teiltöne haben, die nahe an den Teiltönen des zuerst gefundenen Notenobjektes liegen. Durch den nur teilweisen Energieabzug können weitere Notenobjekte in den folgenden Iterationen noch gefunden werden.

[0045] Das Reduzieren der Energie an den gegebenen Frequenzorten in der Energielandschaft $E(f, t, E)$ geschieht vorteilhaft wieder in Form einer in Frequenzrichtung nach oben und unten ausblendenden Fensterfunktion, deren Breite in der Größenordnung von einem Halbton liegt. Falls ein Modell des Obertonspektrums des Klanges bekannt ist, z.B. weil zu dem den Klang erzeugenden Instrument ein Referenzspektrum vorliegt oder modellhaft bekannt ist, kann der Energieentzug in den Obertönen diesem Referenzspektrum oder Modell entsprechend erfolgen. Das Notenobjekt "merkt sich" den durch sie entzogenen Energieanteil für ihre spätere Bewertung, indem ihr dieser Energieanteil zugeschrieben wird.

[0046] In dem Zeitbereich, der durch das neu gefundene Notenobjekt betroffen ist, wird die Relevanzlandschaft wie weiter oben beschrieben neu berechnet, da sich in diesem Zeitbereich die Energielandschaft als Grundlage für die Relevanzlandschaft durch das Entziehen der Energie geändert hat.

[0047] Für das neu gefundene Notenobjekt wird ermittelt, ob es sich in der Zeit und im Verlauf seiner Grundfrequenz mit einem anderen vorher bereits gefundenen Notenobjekt überschneidet. Falls es sich mit einem solchen so überschneidet oder es direkt daran anschließt, dass es plausibel ist, dass es sich um dasselbe Notenobjekt handelt, wird es diesem zugeschlagen (ggf. unter Verlängerung dessen Tonhöhenverlaufes). Anderenfalls wird es als neues Notenobjekt in die Menge der gefundenen Notenobjekte aufgenommen. Da in jedem Iterationsschritt nur z.B. 50% der Energie für das Notenobjekt entzogen werden, wird in der Regel jedes Notenobjekt im Laufe der Iteration mehrfach gefunden.

[0048] Die Iteration wird fortgesetzt, indem in der veränderten Relevanzlandschaft wieder der höchste Punkt gesucht wird. Die Iteration wird bis zum Erreichen eines Abbruchkriteriums fortgesetzt. Ein vorteilhaftes Abbruchkriterium für die Iteration ist die Reduktion der Energie im Verhältnis zur ursprünglichen Energie in der Energielandschaft. Die Iteration kann z.B. abgebrochen werden, wenn nur noch 10% der ursprünglichen Energie in der Energielandschaft vorhanden ist. Auch dies kann

dem Anwender als zu verändernder Parameter zur Verfügung gestellt werden.

[0049] Das Auffinden der ereignishaften Objekte, die sich durch ein plötzliches Ansteigen des geräuschhaften Anteils im Signal auszeichnen, kann entweder im Zeitbereichssignal erfolgen, indem das Ansteigen vor allem hochpassgefilterter Signalanteile verfolgt wird, oder im Frequenzbereich mit Hilfe der Geräuschhaftigkeitswerte der Bins, die dazu, gewichtet mit den Energien der jeweiligen Bins, für jede Zeitscheibe aufaddiert werden. In beiden Fällen erhält man eine Verlaufskurve des geräuschhaften Anteils im Gesamtsignal. An den Punkten der größten Steigungen dieser Kurve, ggf. durch einen Schwellenwert für die Steigung definiert, sind die ereignishaften Objekte anzunehmen.

[0050] Die im vorhergehenden Schritt gefundenen ereignishaften Objekte können entweder isoliert für sich selbst im Signal vorkommen, wie es bei rein perkussiven Ereignissen der Fall sein wird, oder sie können die Anschlaggeräusche der Notenobjekte sein, die zuvor in der Iteration gefunden worden sind, wie es bei gezupften oder angeschlagenen tonalen Instrumenten wie Gitarre, Klavier etc. der Fall sein wird. Um das zu unterscheiden, wird für jedes gefundene ereignishaftige Objekt am Zeitpunkt seines Auftretens untersucht, ob an einer oder an mehreren der dort befindlichen Notenobjekte unmittelbar nach dem Auftreten des Ereignisses ein signifikanter Anstieg ihrer Energie auftritt. Falls dies der Fall ist, wird das ereignishaftige Objekt als Anschlag des Notenobjektes verstanden und diesem zugeordnet. Falls der Energieanstieg bei mehreren Noten der Fall ist, wird das ereignishaftige Objekt allen diesen Noten zugeordnet. Falls der Energieanstieg mitten in einem Notenobjekt auftritt, wird das Notenobjekt dort getrennt und ab dort als neues Notenobjekt verstanden. Wenn zum Zeitpunkt des ereignishaften Objekts kein korrespondierendes Notenobjekt gefunden wird, wird das ereignishaftige Objekt als isoliertes perkussives Ereignis verstanden. Figur 7 zeigt die im vorliegenden Beispiel gefundenen Notenobjekte zusammen mit den als senkrechte Striche gekennzeichneten ereignishaften Objekten, die diesen Notenobjekten zugeordnet werden konnten.

[0051] Dem Auffinden der notenhaften Objekte sollte mit Vorteil ein Bewertungsschritt folgen. Bei der Suche nach Notenobjekten in der beschriebenen Iteration werden in der Regel mehr Objekte gefunden, als musikalisch plausibel vorhanden sind. Deshalb wird die Menge der gefundenen Notenobjekte am Ende noch nach verschiedenen Plausibilitätskriterien überprüft und ggf. nicht ausreichend plausible Notenobjekte entfernt. Ein Plausibilitätskriterium ist z.B. die relative Energie und die Verdeckung. Im Allgemeinen werden bei dem geschilderten iterativen Vorgehen zu viele kleine Notenobjekte mit zu wenig Energie gefunden. Deshalb wird untersucht, wie viel Energie die Note im Verhältnis zur gesamten Energie in ihrem Zeitbereich hat. Wenn sie zuwenig relative Energie besitzt, kann sie entfernt werden.

[0052] Manchmal werden auch Objekte als eigenstän-

dige Noten identifiziert, die eigentlich Obertöne einer anderen vorhandenen Note sind. Hier kann z.B. untersucht werden, ob die höhere Note einen eigenständigen Verlauf in Tonhöhe, Amplitude und Dauer aufweist, oder ob sie in diesen Parametern im Wesentlichen wie eine tiefere Note verläuft. Wenn letzteres der Fall ist, kann das Objekt entfernt werden oder der tieferen Note zugeschlagen werden.

[0053] Weitere Bewertungen können nach musikalischen Gesichtspunkten erfolgen. Wenn z.B. ein Notenobjekt in ihrer Tonhöhegegend sehr isoliert steht (sehr hoch oder sehr tief, wenn sich dort keine anderen Noten befinden), ist sie musikalisch unwahrscheinlich. Wenn z.B. eine Note sich mit anderen Noten in ihrer tonhöhen- und zeitlichen Nachbarschaft zu einer aufsteigenden oder absteigenden Linie verbindet, ist sie musikalisch sehr wahrscheinlich, auch wenn sie sonst eher schwach ist, usw. All diese Kriterien lassen sich mathematisch abbilden und z.B. gewichten, um zu einer möglichst plausiblen Menge von Notenobjekte zu gelangen.

[0054] Diesem geschilderten Identifizierungsschritt kann auch noch ein Eingreifen durch den Anwender folgen, dem die aufgefundenen Notenobjekte in geeigneter Weise graphisch dargestellt werden, z.B. in der in Fig. 7 dargestellten Art, und der z.B. maus- und/oder menügesteuert als eine Note identifizierte Objekte noch aufteilen kann oder getrennte Noten zu einem Objekt vereinen kann. Der Anwender kann natürlich auch einzelne Objekte löschen oder weitere zu berücksichtigende Objekte hinzufügen. Dazu können ihm die Objekte zur Aktivierung angeboten werden, die zuvor bei der automatischen Bewertung als zu wenig relevant bewertet wurden.

[0055] Die automatische Identifizierung kann optional dadurch optimiert werden, dass die Noten des aufgenommenen Musikstück hinterlegt werden, so dass in dem oben genannten Verfahren anhand der hinterlegten Noten gezielt versucht wird, dort Grundtöne aufzufinden, die den Frequenzen der hinterlegten Noten entsprechen. Dies kann z.B. geschehen durch Auswertung einer hinterlegten MIDI-Datei, die die Noten der aufgenommenen Komposition enthält. Es können alternativ auch gleichzeitig mit der Aufnahme des eigentlich verwendeten Gesamtsignals Stützspuren aufgenommen werden, wie z.B. durch Einzelmikrofonierung der beteiligten Instrumentalisten oder Sänger, oder bei einer Gitarre durch Einzelabnahme der Saiten. In dem dann ausgeprägter monophonen Signal der Einzelstimmen können die gewünschten Noten des Gesamtsignals eindeutiger identifiziert und somit der Gesamtklang besser zur Bearbeitung separiert werden.

b) Klangzuordnung zu Notenobjekten

[0056] Nachdem die Identifizierung der am Gesamtklang beteiligten Einzelobjekte abgeschlossen ist, kann in einem nächsten Schritt die klangliche Zerlegung des Gesamtklanges erfolgen. Dabei ist die möglichst genaue Bestimmung der vorhandenen Notenobjekte, ihres Ton-

höhen- und Amplitudenverlaufes und des Charakters ihres Einsetzens eine Einflussgröße für die Qualität des Ergebnisses der klanglichen Zerlegung. Bei der nachfolgend beschriebenen klanglichen Zerlegung wird der Gesamtklang nur so weit in einzelne klangliche Objekte zerlegt, wie es zur Resynthese des neuen Gesamtklanges notwendig ist. Wenn also z.B. in einem gefundenen komplexen Akkord vom Anwender nur eine einzelne Note angefasst und in ihrer Tonhöhe verändert wird, so muss auch nur das Signal dieser einen Note extrahiert und von dem ursprünglichen Signal subtrahiert werden. Es werden also um so mehr einzelne Signale erzeugt, je mehr Noten verändert werden. Jedes ist dann jeweils ein monophon-periodisches Signal und kann mit bereits bekannten Verfahren zur zeit- und tonhöhenunabhängigen Wiedergabe verändert, summiert und abgespielt werden.

[0057] In einem ersten Schritt der Klangzerlegung werden die ereignishaften Objekte aus dem Originalsignal extrahiert. Wenn das Originalsignal in zu den einzelnen Notenobjekten gehörende Einzelsignale zerlegt wird, werden dabei durch die Aufteilung des Frequenzspektrums die ereignishaften Orte im Zeitsignal verschmiert. Deshalb ist es sinnvoll, zunächst die Orte der ereignishaften Objekte aus dem Zeitsignal zu separieren und die Zerlegung in die Einzelsignale der Notenobjekte an dem so entstehenden Restsignal durchzuführen. Gleichwohl kann dieser optionale Verfahrensschritt auch weggelassen werden.

[0058] Zur Separation der ereignishaften Objekte wird zunächst ein modifiziertes Zeitsignal erzeugt, in dem die klanghaften Anteile möglichst vollständig abgezogen sind. Dazu werden im Frequenzbereich die Magnituden aller Bins mit den Geräuschhaftigkeitswerten der Bins multipliziert und mit diesen neuen Magnituden und den Originalphasen über die FFT wieder ein Zeitsignal erzeugt. Optional können in die Faktoren für die Magnituden noch Faktoren zur Dämpfung der tieferen Signalanteile eingehen, da für die ereignishaften Objekte oft die höheren Anteile relevanter sind. Aus diesem neuen geräuschhaften Zeitsignal wird an den Stellen, an denen in dem weiter oben beschriebenen Identifizierungsschritt ereignishaftige Objekte gefunden worden sind, das Zeitsignal der ereignishaften Objekte mit einer geeigneten Fensterfunktion herausgetrennt, z.B. mit einer kurzen Anstiegszeit von etwa 0.005 s und einer Ausklingzeit von etwa 0.05 s. Diese kurzen Zeitsignale der ereignishaften Objekte werden vom Original-Zeitsignal abgezogen und z.B. separat gespeichert.

[0059] Es folgt dann eine Separation der Notenobjekte aus dem Originalsignal. Die Aufteilung des Originalsignals (von dem die ereignishaften Anteile abgezogen wurden) in die Einzelklänge der Notenobjekte geschieht im Frequenzbereich. Dazu wird das Originalsignal nach seiner Modifizierung durch die Separation der Ereignisobjekte (siehe oben) zunächst neu in den Frequenzbereich transformiert.

[0060] Die Aufteilung in die Einzelklänge der Noten-

objekte beruht darauf, dass jedes Notenobjekt in jeder Zeitscheibe einen "Anspruch" auf einen spektralen Anteil des Gesamtsignals $F(f, t, E)$ anmeldet. Dieser Anspruch wird mathematisch dargestellt durch spektrale Anteilsfaktoren, die errechnet werden für jedes Notenobjekt aus einer spektralen Anteilsfunktion, welches z.B. aus einem Modell des Klanges einer einzelnen Note gewonnen wird. Dieses Modell kann einfach vorgegeben werden, oder es kann dem realen Klang eines Instrumentes nachgebildet sein, wenn das Instrument bekannt ist, auf das das Notenobjekt zurückgeht. Das Modell beruht im vorliegenden Beispiel auf den folgenden Komponenten: Es wird angenommen, dass zum Grundtonverlauf des Notenobjektes Obertöne als spektrale Komponenten in ganzzahligen Vielfachen der Grundtonfrequenz existieren. Es wird weiter angenommen, dass die Amplituden der Obertöne zu einem Grundton einem Obertonmodell folgen. Das kann im einfachsten Fall die Abnahme der Amplitude mit dem Kehrwert der Obertonnummer sein. Das Obertonmodell kann aber auch den Obertonamplitudenverlauf abbilden, der aus einem empirischen Klang abgeleitet wurde. Es wird schließlich angenommen, dass die Obertonamplituden in einem Verhältnis zum Verlauf der Grundtonenergie des Notenobjektes stehen. Dies kann im einfachsten Fall als proportional zur Grundtonenergie angenommen werden, es kann aber auch ein anderer Zusammenhang aus einem empirischen Klang abgeleitet werden.

[0061] Basierend auf diesen Annahmen wird eine spektrale Anteilsfunktion vorgegeben, die z.B. für unterschiedliche Instrumente unterschiedlich sein kann, und es werden für jedes Notenobjekt in jeder Zeitscheibe die spektralen Anteilsfaktoren errechnet, also die Ansprüche, die dieses Objekt an jedem Bin hat.

[0062] Wenn mehrere Anteilsfunktionen hinterlegt sind, kann z.B. vorgesehen sein, daß der Anwender eine der Anteilsfunktionen auswählt. Es kann auch eine automatische Auswahl erfolgen, z.B. wenn der Anwender eingibt, mit welchem Instrument das Notenobjekt gespielt wurde, oder wenn automatisch erkannt wird, daß das Notenobjekt auf einem bestimmten Instrument gespielt wurde, was sich z.B. daraus ergeben kann, daß die Obertonamplitudenverhältnisse eines Notenobjektes einer hinterlegten Anteilsfunktion entsprechen.

[0063] Die Höhe der errechneten Anteilsfaktoren hängt ab von den Teiltonfrequenzen und den Amplituden der Teiltöne, die z.B. aus dem zugrunde gelegten Modell des Klanges einer einzelnen Note folgen. Weiter hängt die Höhe der Anteilsfaktoren ab von der Nähe oder Entfernung der Teiltonfrequenz zur Augenblicksfrequenz des jeweiligen Bins. Die Höhe der Anteilsfaktoren in Abhängigkeit von der Entfernung kann z.B. über eine Gewichtungskurve im Frequenzbereich eingehen, die z.B. breit genug sein sollte, dass auch leichte Abweichungen von der gegebenen Frequenz zugelassen werden. Andererseits sollte die Gewichtungskurve im Zentralbereich so schmal sein, dass die Obertonanteilsfaktoren von unterschiedlichen gleichzeitig erklingenden Noten mit un-

terschiedlicher Grundtonhöhe ausreichend gut getrennt und die Obertöne der richtigen Note zugewiesen werden können. Eine geeignete Gewichtungskurve zur Bewertung der Entfernung der Frequenzen kann z.B. ein zur vierten Potenz erhobenes von-Hann-Fenster sein, das in seiner Gesamtbreite z.B. zwei Halbtönen entspricht.

[0064] Wenn in der jeweils aktuellen Zeitscheibe alle dort als klingend gefundenen Notenobjekte ihre Ansprüche über die Anteilsfaktoren an allen Bins angemeldet haben, wird die Summe der Anteilsfaktoren aller Noten für je ein Bin auf 1 normiert. Für jedes Notenobjekt wird ein eigenes Zeitsignal mit der Dauer dieses Notenobjektes angelegt. Für jede Zeitscheibe werden die Magnituden oder ein anderer geeigneter energiehafter Wert aller Bins gemäß den normierten Anteilsfaktoren auf die Notenobjekte aufgeteilt. Diese Anteile der einzelnen Notenobjekte im Frequenzbereich werden mit den Originalphasen über die FFT in den Zeitbereich zurücktransformiert und die Zeitsignale auf die einzelnen Zeitsignale der Notenobjekte akkumuliert.

[0065] Da die Magnitudenanteile oder andere energiehafte Anteile zuvor verändert wurden, sind die Enden des Signals in den Zeitfenstern nach Rücktransformation in den Zeitbereich nicht mehr auf 0 ausgeblendet, was zu unerwünschten Artefakten führt. Deshalb sollte das Ergebnis der Rücktransformation nochmals gefenstert werden. Dazu ist es zweckmäßig, aus den Werten der eigentlich vorgesehenen Fensterfunktion die Wurzel zu ziehen und dieses Fenster dann vor der FFT und nach der inversen FFT anzuwenden.

[0066] Schließlich werden die Notenobjekte mit den ereignishaften Anteilen zusammengeführt. Weiter oben war bereits beschrieben worden, dass den Notenobjekten die ereignishaften Objekte zugeordnet wurden, und dass für die ereignishaften Objekte ein Zeitsignal erzeugt wurde. Jetzt kann den Notenobjekten, denen ereignishaften Objekte zugeordnet wurden, dieses Zeitsignal an ihren Anfang aufaddiert werden. Wenn einem ereignishaften Objekt mehrere Notenobjekte zugeordnet worden sind, weil angenommen wurde, dass diese gleichzeitig angeschlagen wurde, wird das Zeitsignal des ereignishaften Objekts in seiner Amplitude auf die zugeordneten Notenobjekte verteilt. Dies kann zweckmäßig im Verhältnis der Energien der Notenobjekte selbst geschehen oder aufgrund des angenommenen Instrumenten-Modells erfolgen..

[0067] Ereignishaft Objekte, denen keine Notenobjekte zugeordnet wurden, können mit ihrem extrahierten Zeitsignal als eigenständige perkussive Objekte zur Verfügung stehen.

[0068] Wenn für alle gefundenen Notenobjekte und den diesen zugeordneten Ereignisobjekten ihr Zeitsignal erzeugt wurde, werden diese Zeitsignale aller Noten vom Originalsignal abgezogen. Weil im der Klangaufteilung zugrunde gelegten Klangmodell vorausgesetzt wurde, dass die Notenobjekte im wesentlichen aus Teiltönen bestehen, die annähernd ganzzahlige Vielfache einer Grundfrequenz sind, wird nicht der gesamte Klang und

damit nicht das gesamte Zeitsignal auf die Einzelobjekte verteilt worden sein. Deshalb verbleibt nach Abzug der Zeitsignale aller Einzelobjekte vom Originalsignal ein Restsignal, das die eher geräuschhaften Anteile enthält. Dieses Restsignal kann bei der Resynthese einfach mit wiedergegeben werden, oder es kann dem Anwender, als ganzes oder zerlegt in weitere Einzelobjekte, zur weiteren Bearbeitung zur Verfügung gestellt werden.

[0069] Die zeitliche Abfolge der oben genannten einzelnen Verfahrensschritte kann auch anders gewählt werden. So kann z.B. die Zuordnung von Ereignisobjekten zu Notenobjekten auch erst unmittelbar vor der Resynthese vorgenommen werden. Das gilt in analoger Weise für andere Verfahrensschritte, wie z.B. die Identifizierung der Ereignis- und Notenobjekte oder das Errechnen von Anteilsfaktoren.

Patentansprüche

1. Verfahren zur klangobjektorientierten Analyse und zur notenobjektorientierten Bearbeitung einer polyphonen, digitalisierten Klangaufnahme, die als Zeitsignal $F(A, t)$ vorliegt, umfassend die folgenden Analyse- und Bearbeitungsschritte:

- abschnittsweises Auslesen des Zeitsignals $F(A, t)$ unter Verwendung einer Fensterfunktion und sich überlappenden Fenster,
- Fourier-Transformation des ausgelesenen Signals in den Frequenzraum, insbesondere durch Ausführen einer diskreten Fourier-Transformation,
- Berechnen eines Energiewertes E zu jedem Bin aus der Frequenzamplitude, die sich aus der Fourier-Transformation ergibt, insbesondere durch Quadrieren der Real- und Imaginärteile oder Bildung eines davon abgeleiteten Energiewertes,
- Erzeugen einer dreidimensionalen Funktion $F(t, f, E)$,
- Identifizierung von Ereignisobjekten,
- Identifizierung von Notenobjekten
- Vergleich des zeitlichen Auftretens von Ereignisobjekten und Notenobjekten und Zuordnung von Ereignisobjekten zu Notenobjekten für den Fall zeitlich plausiblen Auftretens
- Berechnen von spektralen Anteilsfaktoren zu jedem Notenobjekt,
- Zuordnen von Signalanteilen des Frequenzsignals $F(f, t, E)$ zu gefundenen Notenobjekten anhand der berechneten Anteilsfaktoren,
- Rücktransformation der einem Notenobjekt zugeordneten Frequenzsignalanteile in ein Zeitsignal,
- graphische Darstellung der Notenobjekte und/oder Ereignisobjekte in einer Zeit/Frequenzdarstellung auf einem Monitor,

- anwendergesteuerte oder automatisierte Bearbeitung eines oder mehrerer Notenobjekte,
 - Speicherung der Zeitsignale bearbeiteter Notenobjekte,
 - Wiedergabe der gespeicherten Zeitsignale bearbeiteter Notenobjekte zusammen mit dem um das einem Notenobjekt zugeordnete Zeitsignal verminderte Zeitsignal.
2. Verfahren nach Anspruch 1, **dadurch gekennzeichnet, dass** aus der Funktion $F(t, f, E)$ eine Funktion $F'(t, f, E')$ errechnet wird mit Energiewerten E' , die die Summe aller Energiewerte E zu einer Zeit t bei einer Grundfrequenz f und allen Vielfachen davon ist.
3. Verfahren nach Anspruch 2, **dadurch gekennzeichnet, daß** die Energiewerte bei den Vielfachen der Grundfrequenz nach Gewichtung mit einem von 1 abweichenden Faktor addiert werden.
4. Verfahren nach Anspruch 2 oder 3, **dadurch gekennzeichnet, dass** zur Identifizierung der Notenobjekte folgende Verfahrensschritte ausgeführt werden:
- Ermitteln eines Energiemaximums in der Funktion $F'(f, t, E')$,
 - Ermitteln eines zu dem Maximum gehörenden zusammenhängenden Wertefeldes,
 - Zuordnung des gefundenen Wertefeldes zu jeweils einem Notenobjekt.
5. Verfahren nach Anspruch 4, **dadurch gekennzeichnet, dass** die Energiewerte E des dem Notenobjekt zugewiesenen Wertefeldes von der Funktion $F(t, f, E)$ subtrahiert werden.
6. Verfahren nach Anspruch 5, **dadurch gekennzeichnet, dass** die Energiewerte E nur in der Höhe $G \cdot E$ subtrahiert werden, wobei für den Faktor G $0 < G < 1$ gilt.
7. Verfahren nach Anspruch 6, **dadurch gekennzeichnet, dass** der Faktor G ein vom Anwender veränderbarer Parameter ist.
8. Verfahren nach Anspruch 5, 6 oder 7, **dadurch gekennzeichnet, dass** die Maximumsuche an einer Funktion fortgesetzt wird, von der die Energiewerte abgezogen sind oder an einer daraus errechneten Funktion.
9. Verfahren nach einem der vorhergehenden Ansprüche, **dadurch gekennzeichnet, dass** die Maximumsuche iterativ bis zum Erreichen eines Abbruchkriteriums ausgeführt wird.
10. Verfahren nach Anspruch 9, **dadurch gekennzeichnet, dass** ein Energiegesamtwert E_{ges} zur Funktion $F(t, f, E)$ errechnet wird und die Iteration abgebrochen wird, sobald ein bestimmter Anteil $H \cdot E_{\text{ges}}$ dieses Gesamtwertes den gefundenen Notenobjekten zugeordnet worden ist, insbesondere wenn dieser Anteil 90% übersteigt.
11. Verfahren nach Anspruch 10, **dadurch gekennzeichnet, dass** der Faktor H ein vom Anwender veränderbarer Parameter ist.
12. Verfahren nach einem der vorhergehenden Ansprüche, **dadurch gekennzeichnet, dass** ein identifiziertes Notenobjekte bei einer anschließenden automatisierten Bereinigung verworfen wird bei Vorliegen eines oder mehrerer der folgenden Kriterien:
- die Energie des gefundenen Notenobjektes ist im Vergleich zur Gesamtenergie sehr gering,
 - der Tonhöhen- und Amplitudenverlauf des Notenobjektes stimmt im Wesentlichen mit dem eines anderen Notenobjektes niedrigerer Frequenz überein, insbesondere wenn die niedrigere Frequenz als Grundfrequenz in Frage kommt,
 - der Frequenzabstand eines Notenobjektes zu den anderen Notenobjekten ist sehr groß.
13. Verfahren nach einem der vorhergehenden Ansprüche, **dadurch gekennzeichnet, dass** in einem Nachbearbeitungsschritt ein Anwender automatisch identifizierte Notenobjekte trennt, verbindet und/oder löscht.
14. Verfahren nach einem der vorhergehenden Ansprüche, **dadurch gekennzeichnet, dass** zu jedem Bin rechnerisch eine Augenblicksfrequenz aus den Phasendifferenzen benachbarter Bins ermittelt wird, die als Frequenz des Bins in der Funktion $F(t, f, E)$ oder $F'(t, f, E')$ verwendet wird.
15. Verfahren nach einem der vorhergehenden Ansprüche, **dadurch gekennzeichnet, dass** zum Auffinden der Ereignisobjekte zu jedem Bin ein Tonalitätswert und/oder Geräuschhaftigkeitswert errechnet wird.
16. Verfahren nach Anspruch 15, **dadurch gekennzeichnet, dass** die Energiewerte jedes Bins mit dem Tonalitätswert gewichtet werden.
17. Verfahren nach einem der vorhergehenden Ansprüche, **dadurch gekennzeichnet, dass** für die Identifizierung von Notenobjekten auf hinterlegte Noten zurückgegriffen wird.
18. Verfahren nach einem der vorhergehenden Ansprüche

che, **dadurch gekennzeichnet, dass** die gefundenen Ereignisobjekte aus dem Zeitsignal $F(A, t)$ extrahiert werden und die Klangzerlegung am verbleibenden Signal erfolgt.

19. Verfahren nach Anspruch 18, **dadurch gekennzeichnet, dass** die Ereignisobjekte separat gespeichert werden.

20. Verfahren nach einem der vorhergehenden Ansprüche, **dadurch gekennzeichnet, dass** die spektralen Anteilsfaktoren eines Notenobjektes aus einer hinterlegten spektralen Anteilsfunktion errechnet werden.

21. Verfahren nach Anspruch 20, **dadurch gekennzeichnet, daß** die hinterlegte spektrale Anteilsfunktion die mathematische Abbildung eines Klangmodells des Notenobjektes ist, das eines oder mehrere der folgenden Postulate aufstellt:

- es gibt zu einem Grundton mit einer Grundtonfrequenz zugehörige spektrale Komponenten auf ganzzahligen Vielfachen der Grundtonfrequenz, nämlich Obertöne auf Obertonfrequenzen,
- der Amplitudenverlauf der Obertöne zu einem Grundton folgt einer Gesetzmäßigkeit, die sich aus einem Obertonmodell oder einem empirisch ermittelten Obertonamplitudenverlauf ergibt,
- die Obertonamplituden stehen in einem festen Verhältnis zum Verlauf der Grundtonenergie.

22. Verfahren nach Anspruch 20, **dadurch gekennzeichnet, daß** die hinterlegte spektrale Anteilsfunktion die mathematische Abbildung des Klanges einer auf einem bestimmten Instrument gespielten Note ist.

23. Verfahren nach Anspruch 20, **dadurch gekennzeichnet, daß** mehrere spektrale Anteilsfunktionen hinterlegt sind, insbesondere für mehrere Instrumente verschiedene Anteilsfunktionen, die insbesondere solche nach Anspruch 22.

24. Verfahren nach Anspruch 23, **dadurch gekennzeichnet, daß** vom Anwender eine der mehreren Anteilsfunktionen ausgewählt wird.

25. Verfahren nach Anspruch 23, **dadurch gekennzeichnet, daß** einem Notenobjekt eine der mehreren Anteilsfunktionen automatisch zugewiesen wird, wenn vom Anwender eingegeben oder automatisiert erkannt wird, auf welchem Instrument die Note gespielt wurde..

26. Verfahren nach einem der Ansprüche 20 bis 25, **dadurch gekennzeichnet, daß** die spektrale Anteils-

funktion über eine Fensterfunktion eine Gewichtung im Frequenzbereich mit einer vorbestimmten Frequenzbreite vornimmt.

27. Verfahren nach einem der vorhergehenden Ansprüche, **dadurch gekennzeichnet, dass** ein Restsignal errechnet wird durch Subtraktion aller den Notenobjekten und Ereignisobjekten zugewiesenen Zeitsignale vom ursprünglichen Zeitsignal.

28. Verfahren nach Anspruch 27, **dadurch gekennzeichnet, dass** das Restsignal einer weiteren Identifizierung von Noten- oder Ereignisobjekten unterzogen wird.

29. Verfahren nach einem der vorhergehenden Ansprüche, **dadurch gekennzeichnet, dass** für die Klangwiedergabe nach Bearbeitung eines Notenobjektes der Klanganteil des Notenobjektes vom Gesamtklang abgezogen wird, und das so erhaltene Differenzsignal zusammen mit dem Klanganteil des Notenobjektes abgespielt wird.

30. Verfahren nach Anspruch 27, **dadurch gekennzeichnet, dass** bei der Klangwiedergabe nach der Bearbeitung eines Notenobjektes durch einen Anwender das Restsignal mit wiedergegeben wird.

31. Computer-Programm mit einem Programmcode zur Durchführung des Verfahrens nach einem der vorhergehenden Ansprüche, wenn das Computer-Programm auf einem Computer abläuft.

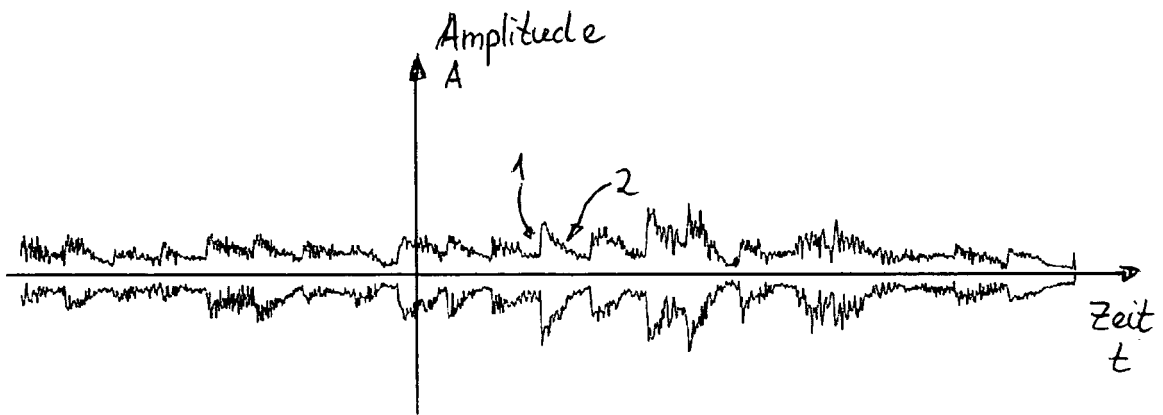


Fig. 1

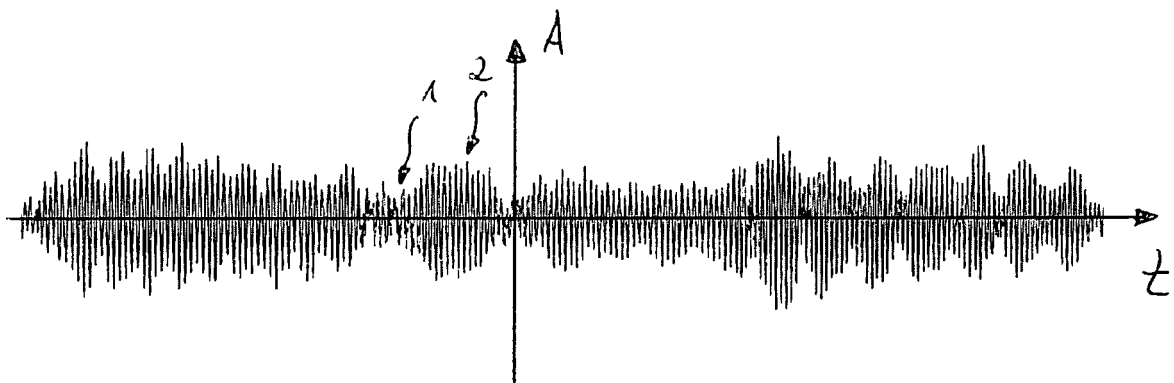


Fig. 2

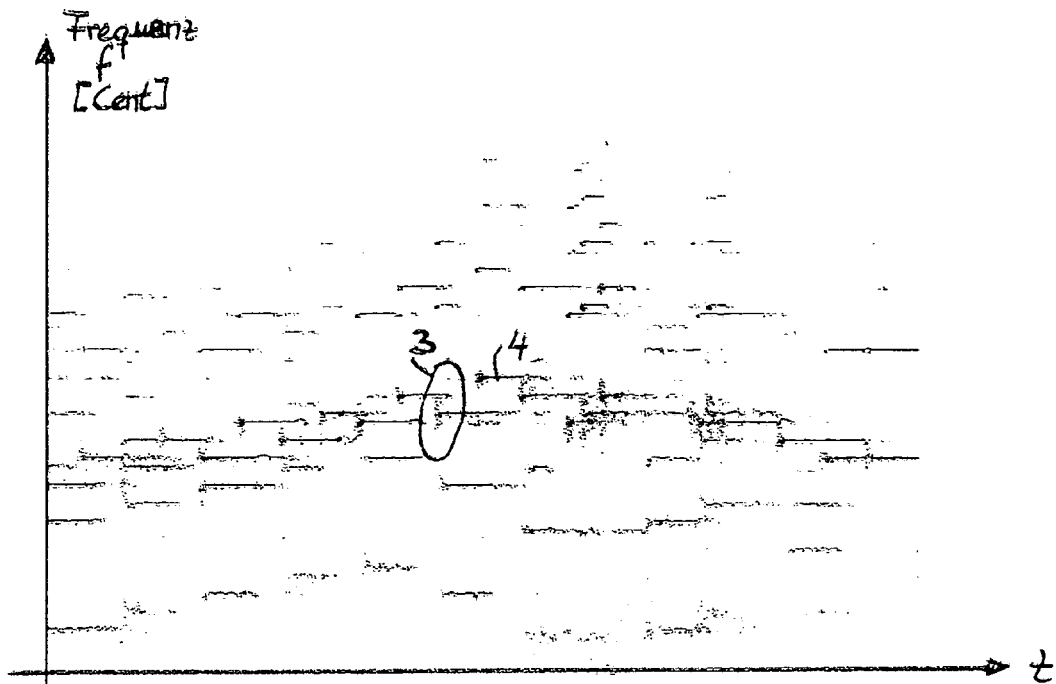


Fig. 3

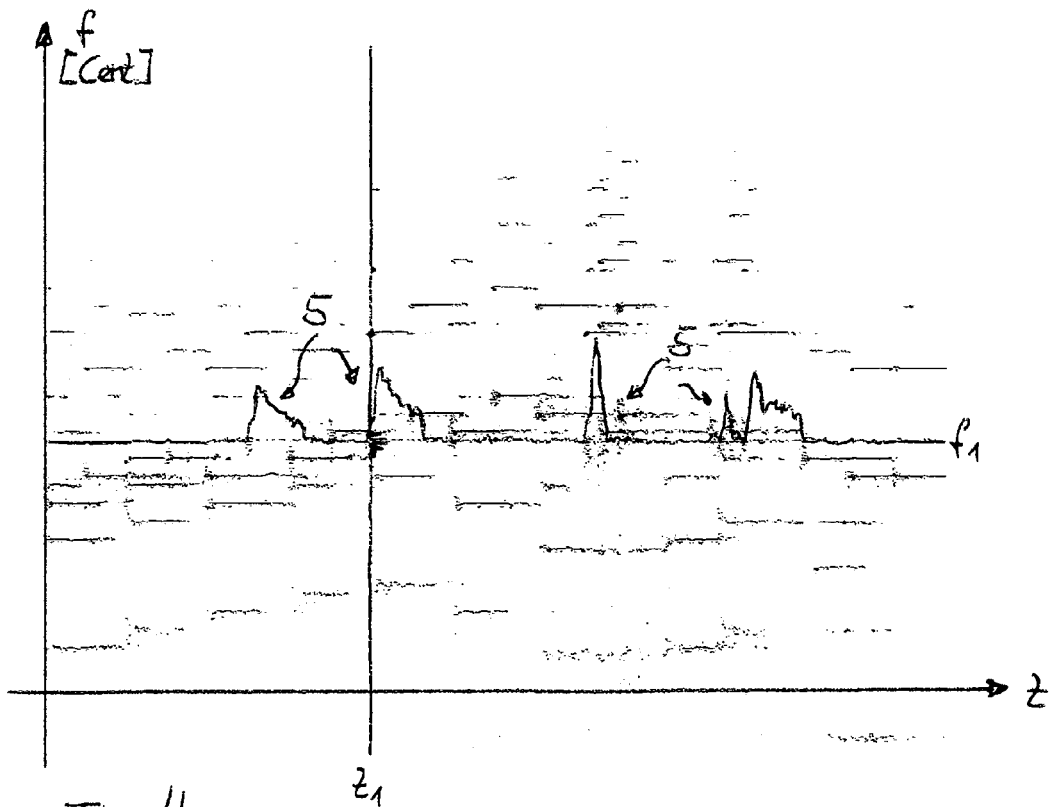
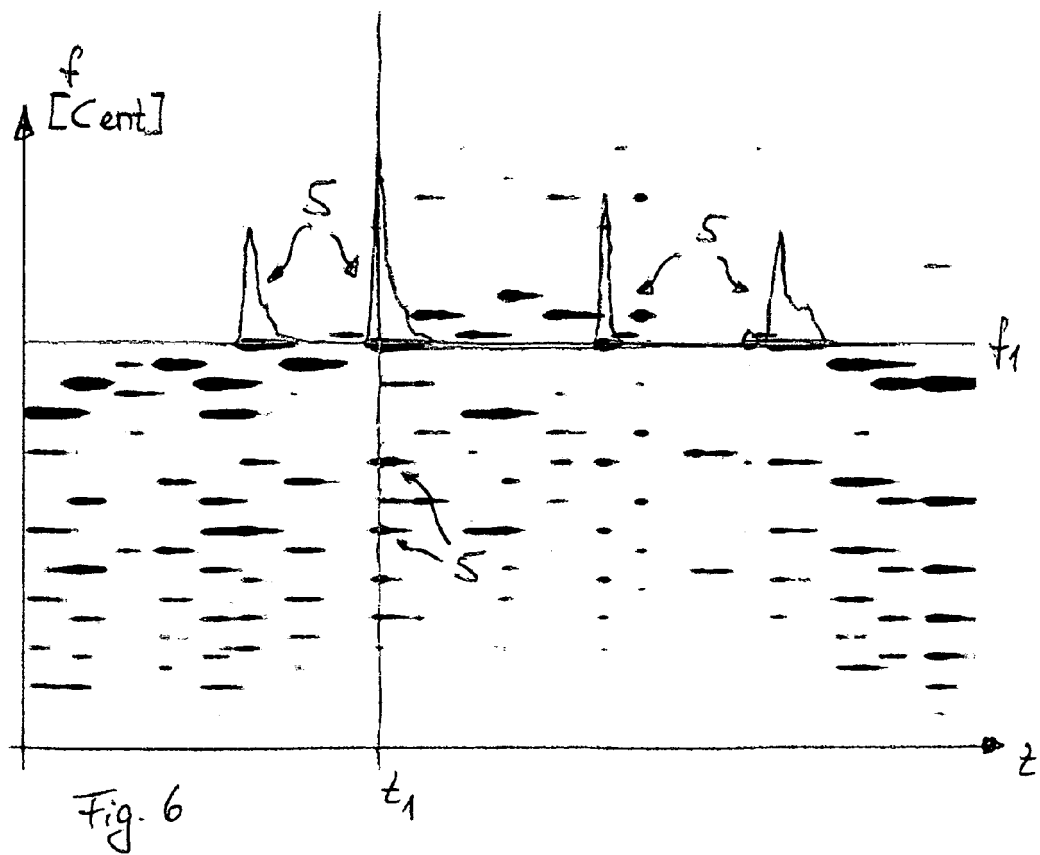
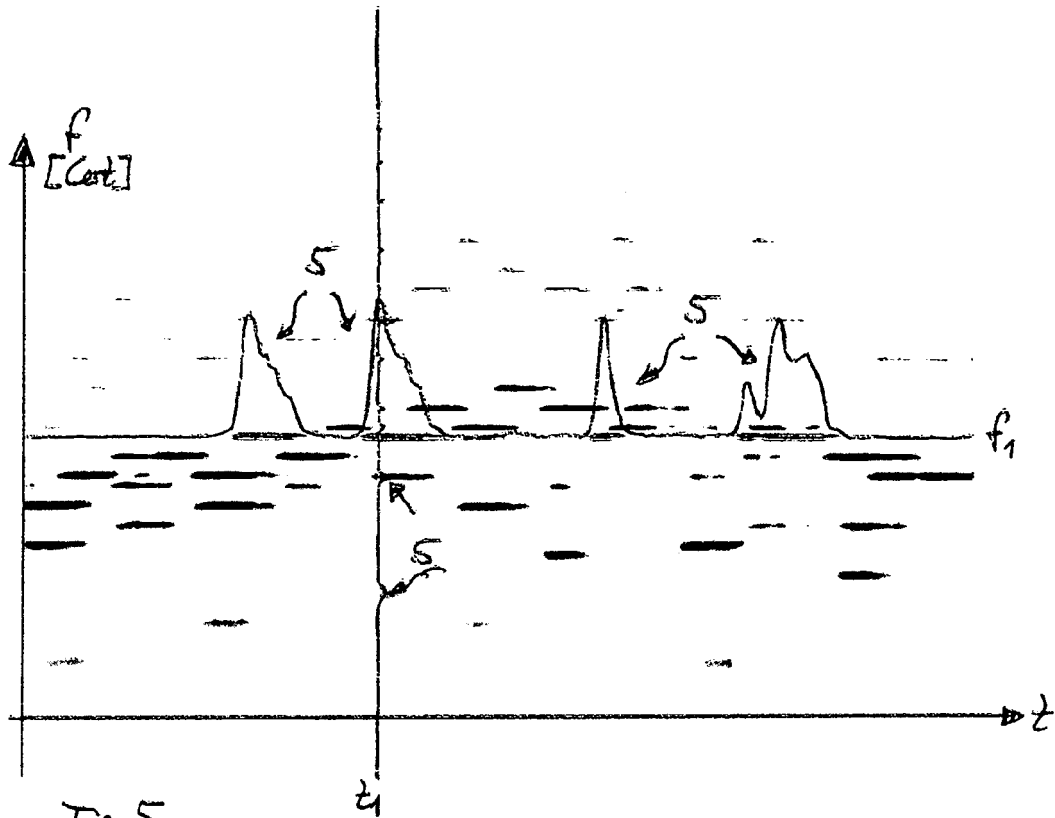


Fig. 4



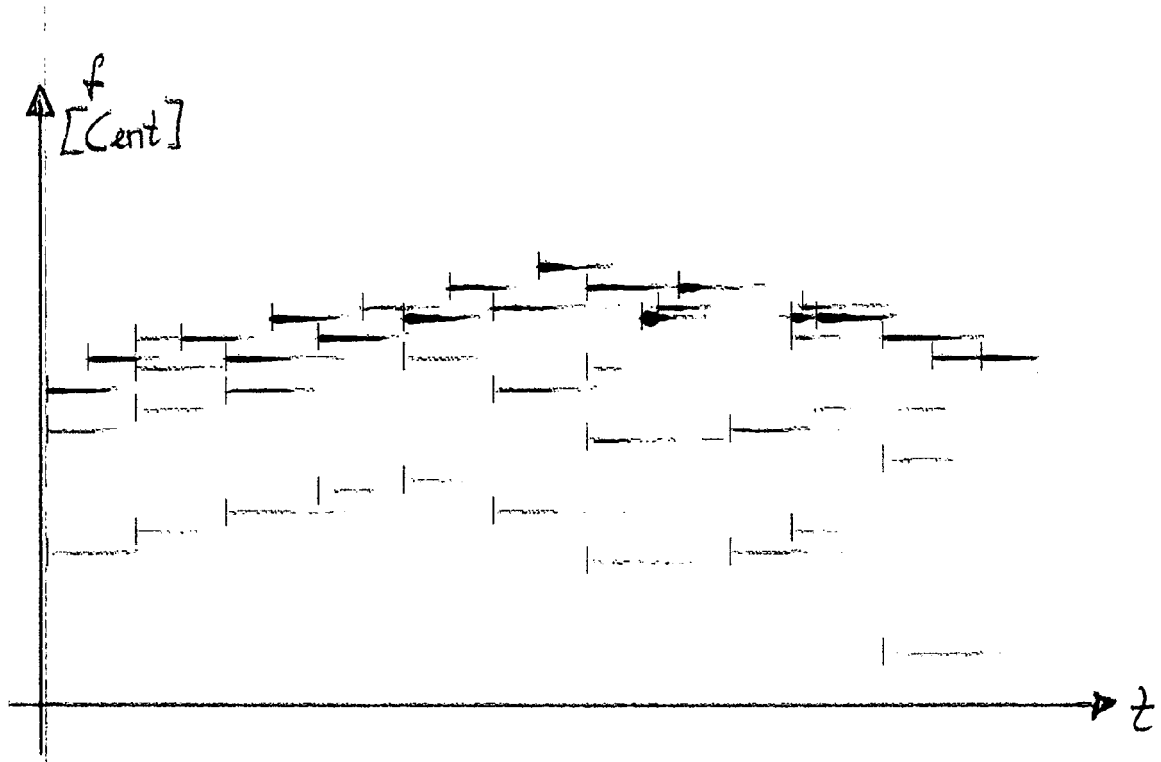


Fig. 7



EUROPÄISCHER RECHERCHENBERICHT

 Nummer der Anmeldung
EP 09 00 1575

EINSCHLÄGIGE DOKUMENTE			
Kategorie	Kennzeichnung des Dokuments mit Angabe, soweit erforderlich, der maßgeblichen Teile	Betrifft Anspruch	KLASSIFIKATION DER ANMELDUNG (IPC)
X	M. EVERY AND AL.: "Separation of synchronous pitched notes by spectral filtering of harmonics" IEEE Transactions on Audio, Speech and Language Processing IEEE USA, Bd. 14, Nr. 5, September 2006 (2006-09), Seiten 1845-1856, XP002533838 ISSN: 1558-7916 * das ganze Dokument *	1-31	INV. G10H1/06
X	RÉMI GRIBONVAL ET AL.: "Harmonic Decomposition of Audio Signals With Matching Pursuit" IEEE TRANSACTIONS ON SIGNAL PROCESSING, IEEE SERVICE CENTER, NEW YORK, NY, US, Bd. 51, Nr. 1, 1. Januar 2003 (2003-01-01), XP011080322 ISSN: 1053-587X Abschnitt II.A, Abschnitte V-VIII * Zusammenfassung; Abbildungen 1-5 *	1-4,20, 21	RECHERCHIERTE SACHGEBIETE (IPC)
X	CHRIS DUXBURY, NICOLAS CHÉTRY, MARK SANDLER, D MIKE DAVIES: "Efficient Two stage implementation of the Harmonic Matching Pursuit" EUSIPCO 2004, [Online] 2004, XP002533839 Gefunden im Internet: URL: http://www.eurasip.org/Proceedings/Eusipco/Eusipco2004/defevent/papers/cr1814.pdf fslVF1g> [gefunden am 2009-06-24] * das ganze Dokument *	1-11, 13-26	G10H
Der vorliegende Recherchenbericht wurde für alle Patentansprüche erstellt			
Recherchenort München		Abschlußdatum der Recherche 24. Juni 2009	Prüfer Lecointe, Michael
KATEGORIE DER GENANNTEN DOKUMENTE X : von besonderer Bedeutung allein betrachtet Y : von besonderer Bedeutung in Verbindung mit einer anderen Veröffentlichung derselben Kategorie A : technologischer Hintergrund O : nichtschriftliche Offenbarung P : Zwischenliteratur		T : der Erfindung zugrunde liegende Theorien oder Grundsätze E : älteres Patentedokument, das jedoch erst am oder nach dem Anmeldedatum veröffentlicht worden ist D : in der Anmeldung angeführtes Dokument L : aus anderen Gründen angeführtes Dokument & : Mitglied der gleichen Patentfamilie, übereinstimmendes Dokument	

 1
EPO FORM 1503 03.82 (P04C03)



EUROPÄISCHER RECHERCHENBERICHT

 Nummer der Anmeldung
EP 09 00 1575

EINSCHLÄGIGE DOKUMENTE			
Kategorie	Kennzeichnung des Dokuments mit Angabe, soweit erforderlich, der maßgeblichen Teile	Betrifft Anspruch	KLASSIFIKATION DER ANMELDUNG (IPC)
X	KLAPURI A P: "Multiple fundamental frequency estimation based on harmonicity and spectral smoothness" IEEE TRANSACTIONS ON SPEECH AND AUDIO PROCESSING, IEEE SERVICE CENTER, NEW YORK, NY, US, Bd. 11, Nr. 6, 1. November 2003 (2003-11-01), Seiten 804-816, XP011104552 ISSN: 1063-6676 * Abschnitt II * * Zusammenfassung; Abbildungen 1-6 *	1,20-26	
A	VIRTANEN T ET AL: "Separation of harmonic sounds using linear models for the overtone series" 2002 IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING. PROCEEDINGS. (ICASSP). ORLANDO, FL, MAY 13 - 17, 2002; [IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING (ICASSP)], NEW YORK, NY : IEEE, US, Bd. 2, 13. Mai 2002 (2002-05-13), Seiten II-1757, XP010804234 ISBN: 978-0-7803-7402-7 Abschnitt 2 * Zusammenfassung; Abbildung 1 *	20-26	
A	US 6 057 502 A (FUJISHIMA TAKUYA [JP]) 2. Mai 2000 (2000-05-02) * Zusammenfassung *	2-4	
Der vorliegende Recherchenbericht wurde für alle Patentansprüche erstellt			
Recherchenort München		Abschlußdatum der Recherche 24. Juni 2009	Prüfer Lecointe, Michael
KATEGORIE DER GENANNTEN DOKUMENTE X : von besonderer Bedeutung allein betrachtet Y : von besonderer Bedeutung in Verbindung mit einer anderen Veröffentlichung derselben Kategorie A : technologischer Hintergrund O : nichtschriftliche Offenbarung P : Zwischenliteratur		T : der Erfindung zugrunde liegende Theorien oder Grundsätze E : älteres Patentdokument, das jedoch erst am oder nach dem Anmeldedatum veröffentlicht worden ist D : in der Anmeldung angeführtes Dokument L : aus anderen Gründen angeführtes Dokument & : Mitglied der gleichen Patentfamilie, übereinstimmendes Dokument	

 1
EPO FORM 1503 03.82 (P04C03)

ANHANG ZUM EUROPÄISCHEN RECHERCHENBERICHT ÜBER DIE EUROPÄISCHE PATENTANMELDUNG NR.

EP 09 00 1575

In diesem Anhang sind die Mitglieder der Patentfamilien der im obengenannten europäischen Recherchenbericht angeführten Patentdokumente angegeben.

Die Angaben über die Familienmitglieder entsprechen dem Stand der Datei des Europäischen Patentamts am 1.1.2010. Diese Angaben dienen nur zur Unterrichtung und erfolgen ohne Gewähr.

24-06-2009

Im Recherchenbericht angeführtes Patentdokument	Datum der Veröffentlichung	Mitglied(er) der Patentfamilie	Datum der Veröffentlichung
US 6057502 A	02-05-2000	JP 3826660 B2	27-09-2006
		JP 2000298475 A	24-10-2000

Für nähere Einzelheiten zu diesem Anhang : siehe Amtsblatt des Europäischen Patentamts, Nr.12/82

IN DER BESCHREIBUNG AUFGEFÜHRTE DOKUMENTE

Diese Liste der vom Anmelder aufgeführten Dokumente wurde ausschließlich zur Information des Lesers aufgenommen und ist nicht Bestandteil des europäischen Patentdokumentes. Sie wurde mit größter Sorgfalt zusammengestellt; das EPA übernimmt jedoch keinerlei Haftung für etwaige Fehler oder Auslassungen.

In der Beschreibung aufgeführte Patentdokumente

- EP 0750776 B1 [0003]
- DE 69614938 T2 [0003]
- DE 102004049477 A1 [0004] [0006]
- WO 02084641 A1 [0004]