



Europäisches Patentamt
European Patent Office
Office européen des brevets



(11) **EP 1 093 109 A1**

(12) **EUROPÄISCHE PATENTANMELDUNG**

(43) Veröffentlichungstag:
18.04.2001 Patentblatt 2001/16

(51) Int. Cl.⁷: **G10H 1/00**

(21) Anmeldenummer: **00120872.7**

(22) Anmeldetag: **25.09.2000**

(84) Benannte Vertragsstaaten:
**AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU
MC NL PT SE**

Benannte Erstreckungsstaaten:
AL LT LV MK RO SI

(30) Priorität: **11.10.1999 DE 19948974**

(71) Anmelder:
**NOKIA MOBILE PHONES LTD.
02150 Espoo (FI)**

(72) Erfinder: **Theimer, Wolfgang
44789 Bochum (DE)**

(74) Vertreter:
**TER MEER STEINMEISTER & PARTNER GbR
Mauerkircherstrasse 45
81679 München (DE)**

(54) **Verfahren zum Erkennen und Auswählen einer Tonfolge, insbesondere eines Musikstücks**

(57) Die Erfindung betrifft Verfahren zum Erkennen und zum Auswählen einer Tonfolge, insbesondere eines Musikstücks, die es einem Benutzer ermöglichen, ein bestimmtes Musikstück anzufordern, in dem er einen Abschnitt des Musikstücks singt, dessen Titel er nicht weiß. Dieses Verfahren zeichnet sich dadurch aus, daß eine zumindest einem Abschnitt der auszuwählenden Tonfolge zumindest teilweise entsprechende Tonfolge eingegeben wird, die Töne der eingegebenen Tonfolge in eine Notenfolge umgesetzt werden, zum Suchen der auszuwählenden Tonfolge deren Notenfolge nacheinander mit entsprechenden Notenfolgen einer Vielzahl von Tonfolgen verglichen wird, um Bezeichnungen einer oder mehrerer Tonfolgen zu ermitteln, deren Notenfolge oder -folgen mit der Notenfolge der auszuwählenden Tonfolge in vorbestimmter Weise übereinstimmt bzw. übereinstimmen, und die ermittelten Bezeichnungen als Liste oder Tonfolgen ausgegeben werden, so daß ein Benutzer anhand der Bezeichnungsliste oder Tonfolge die gewünschte Tonfolge auswählen kann.

EP 1 093 109 A1

Beschreibung

[0001] Die Erfindung betrifft sowohl ein Verfahren zum Erkennen als auch zum Auswählen einer Tonfolge, insbesondere eines Musikstücks.

[0002] Die heutigen Multimedadienste ermöglichen es ihren Benutzern, Musikstücke, Videoclips und auch graphische Informationen auf entsprechende Anforderung aus entsprechenden Datenbanken abzurufen, um die gewünschten Musikstücke oder dergleichen wiedergeben und/oder speichern zu können. Bei weitersteigenden Datenübertragungsgeschwindigkeiten und sinkenden Speicherplatzkosten werden in Zukunft auch Filme von entsprechenden Anbietern abrufbar sein.

[0003] Beispielsweise ist es derzeit im Internet möglich, daß ein Benutzer sich von einem entsprechenden Anbieter Musiktitel oder dergleichen übertragen läßt, die dann entweder in einer benutzerseitigen Datenbank gespeichert oder zum Herstellen einer CD verwendet werden. Eine derartige Anforderung von Musikstücken oder dergleichen ist jedoch auch über Mobilfunkdienste möglich.

[0004] Um einen bestimmten Musiktitel zu erhalten, muß der Benutzer den Namen oder die Bezeichnung des Musikstücks eingeben und an den entsprechenden Diensteanbieter übertragen. Die Musiktiteldatenbank des Diensteanbieters wird dann auf das angeforderte Musikstück hin durchsucht um dieses, falls es in der Datenbank verfügbar ist, zum anfordernden Benutzer zu übertragen.

[0005] Um einem Benutzer auch dann einen gewünschten Musiktitel liefern zu können, wenn dieser die Bezeichnung des Musikstücks nicht ganz genau kennt, werden bei der Suche in der Datenbank des Diensteanbieters auch assoziative Suchalgorithmen verwendet, die trotz geringer Abweichungen der eingegebenen Bezeichnung vom tatsächlichen Namen des Musikstücks dieses identifizieren können oder zumindest eine Auswahl von mehreren Musikstücken mit ähnlichen Bezeichnungen anbieten.

[0006] Möchte jedoch ein Benutzer ein Musikstück anfordern, das ihm sehr gut gefällt, dessen Titel er jedoch nicht oder allenfalls nur sehr vage kennt, so ist es für ihn derzeit praktisch unmöglich dieses Musikstück anzufordern.

[0007] Davon ausgehend liegt der Erfindung die Aufgabe zugrunde, Verfahren zum Erkennen und zum Auswählen einer Tonfolge, insbesondere eines Musikstückes bereitzustellen, die es einem Benutzer ermöglichen, eine Tonfolge oder ein Musikstück aufzufinden und auszuwählen, dessen Bezeichnung er nicht kennt.

[0008] Diese Aufgabe wird hinsichtlich des Erkennens einer Tonfolge durch das Verfahren nach Anspruch 1 und hinsichtlich des Auswählens einer Tonfolge durch das Verfahren nach Anspruch 2 gelöst. Vorteilhafte Ausgestaltungen und Weiterbildungen der Erfindung sind in den Unteransprüchen beschrieben.

[0009] Erfindungsgemäß werden also zum Erkennen einer Tonfolge zunächst die Töne der zu erkennenden Tonfolge in eine Notenfolge umgesetzt, anschließend wird zum Suchen der zu erkennenden Tonfolge deren Notenfolge nacheinander mit entsprechenden Notenfolgen einer Vielzahl von Tonfolgen verglichen und dann werden Bezeichnungen der Tonfolge oder -folgen ausgegeben, deren Notenfolge bzw. -folgen mit der Notenfolge der zu erkennenden Tonfolge in vorbestimmter Weise übereinstimmt bzw. übereinstimmen.

[0010] Das erfindungsgemäße Verfahren zum Auswählen einer Tonfolge nutzt dieses Erkennungsverfahren und zeichnet sich dadurch aus, daß eine zumindest einem Abschnitt der auszuwählenden Tonfolge zumindest teilweise entsprechende Tonfolge eingegeben wird, die Töne der eingegebenen Tonfolge in eine Notenfolge umgesetzt werden, zum Suchen der auszuwählenden Tonfolge deren Notenfolge nacheinander mit entsprechenden Notenfolgen einer Vielzahl von Tonfolgen verglichen wird, um Bezeichnungen einer oder mehrerer Tonfolgen zu ermitteln, deren Notenfolge oder -folgen mit der Notenfolge der auszuwählenden Tonfolge in vorbestimmter Weise übereinstimmt, und die ermittelten Bezeichnungen als Liste ausgegeben werden, so daß ein Benutzer anhand der Bezeichnungsliste die gewünschte Tonfolge auswählen kann.

[0011] Der Grundgedanke der vorliegenden Erfindung liegt also darin, daß eine Tonfolge, wie sie sich akustisch dem Benutzer darbietet und von diesem mehr oder weniger genau reproduziert werden kann, zunächst in eine Notenfolge, also in eine Darstellung umgesetzt wird, wie sie beispielsweise auch zum Niederschreiben von Musikstücken verwendet wird, und diese Darstellung der gewünschten Tonfolge mit entsprechenden Notenfolgen, die einzelnen Musikstücken in einer Datenbank eines Diensteanbieters zugeordnet sind, zu vergleichen, so daß der Grad der Übereinstimmung zwischen der eingegebenen, gewünschten Tonfolge und den Musikstücken ermittelt werden kann, um dann die Bezeichnungen der Tonfolge oder -folgen auszugeben, die mit der gewünschten oder auszuwählenden Tonfolge in vorbestimmter Weise übereinstimmen.

[0012] Die Erfindung ermöglicht es somit einem Benutzer, auch Tonfolgen, insbesondere Musikstücke, Videoclips und gegebenenfalls auch Filme über deren Soundtrack anzufordern, wenn ihm nur deren Melodie bekannt ist. Das erfindungsgemäße Verfahren ermöglicht somit eine intuitive Suche in Datenbanken von Musikstücken oder dergleichen, und erleichtert damit deren Nutzung.

[0013] Bei einer ersten Ausgestaltung der Erfindung ist vorgesehen, daß die in einem Benutzerendgerät eingegebene der auszuwählenden Tonfolge entsprechende Tonfolge zu einer Datenbankstation übertragen wird, in der die Liste der Bezeichnungen einer oder mehrerer der auszuwählenden Tonfolge ähnlichen Tonfolgen ermittelt wird, und die

Bezeichnungsliste zur Ausgabe an das Benutzerendgerät übertragen wird.

[0014] Wird als Benutzerendgerät beispielsweise ein Mobiltelefon benutzt, um über Funkkanäle ein bestimmtes Musikstück bei einem Diensteanbieter auszuwählen, so ist es insbesondere im Hinblick auf eine gute Ausnutzung der Übertragungsstrecke vorteilhaft, wenn die in ein Benutzerendgerät eingegebene, der auszuwählenden Tonfolge entsprechende Tonfolge im Benutzerendgerät in eine Notenfolge umgesetzt wird, die Notenfolge zu einer Datenbankstation übertragen wird, in der die Liste der Bezeichnungen einer oder mehrerer der auszuwählenden Tonfolge ähnlichen Tonfolgen ermittelt wird, und die Bezeichnungsliste zur Ausgabe an das Benutzerendgerät übertragen wird.

[0015] Um es dem Benutzer zu ermöglichen, auch ein Musikstück, dessen Titel er überhaupt nicht kennt, auszuwählen zu können, ist bei einer besonders vorteilhaften Ausgestaltung der Erfindung vorgesehen, daß zusammen mit jeder Bezeichnung ein für die jeweilige Tonfolge charakteristischer kurzer Ausschnitt der Tonfolge zur Ausgabe an das Benutzerendgerät übertragen wird. Dem Benutzer wird somit nicht nur die Bezeichnung der jeweiligen Tonfolge also der oder die Titel des erkannten Musikstücks oder in Frage kommender Musikstücke angeboten, sondern es wird ihm auch ermöglicht, einen kurzen charakteristischen Auszug aus dem Musikstück, zum Beispiel das Grundthema oder den Refrain abzuhören, so daß er aufgrund der wiedergegebenen charakteristischen Tonfolge seine Auswahl treffen kann.

[0016] Besonders zweckmäßig ist es, wenn bei dem erfindungsgemäßen Verfahren die Tonfolge zur Eingabe in das Benutzerendgerät vom Benutzer gesungen wird.

[0017] Eine besonders vorteilhafte Ausgestaltung des erfindungsgemäßen Verfahrens zeichnet sich dadurch aus, daß zum Umsetzen einer Tonfolge in eine Notenfolge für jeden Ton der Tonfolge die Tonhöhenfrequenz f_p' und die Tondauer d' ermittelt wird, und jedem Ton eine musikalische Note entsprechend seiner Tonhöhenfrequenz f_p und eine musikalisch quantisierte Notendauer d entsprechend einer Tondauerverteilung der Tonfolge zugeordnet wird.

[0018] Hierbei ist es zweckmäßig, wenn zur Festlegung der Notendauer der Töne zunächst der Median der Tondauerverteilung ermittelt wird, und die Tondauer des Medians mit der Notendauer einer 1/4-Note gleichgesetzt wird, sowie jedem Ton durch Vergleich seiner Tondauer mit der ermittelten Notendauer einer 1/4-Note eine entsprechende musikalische quantisierte Notendauer zugeordnet wird.

[0019] Erfindungsgemäß wird also aus dem zeitlichen Verlauf der Tonhöhenfrequenz der jeweilige musikalische Ton bzw. die Note, also beispielsweise C, D, E, F, G, A, H und die Notendauer d ermittelt. Da insbesondere beim Singen der gewünschten Tonfolge die Notendauer d nicht absolut meßbar ist, wird aus der Tondauerverteilung der Median ermittelt und der Notendauer einer 1/4-Note gleichgesetzt. Davon ausgehend lassen sich dann Tondauerintervalle angeben, denen dann die anderen gebräuchlichen Notendauern, also insbesondere 1/32, 1/16, 1/8, 1/2 und 1 zugeordnet werden können.

[0020] Für die Durchführung des Vergleichs zur Feststellung eines Übereinstimmungsgrades in einer Datenverarbeitungsanlage ist es besonders zweckmäßig, wenn jede Tonfolge durch einen Tonhöhenvektor h , der sich aus den einzelnen Noten oder musikalischen Tönen zusammensetzt, und einen Tondauervektor d dargestellt wird, der sich aus den musikalischen quantisierten Notendauern d der einzelnen Töne zusammensetzt.

[0021] Um die Notenfolge einer eingegebenen Tonfolge auch dann erfolgreich mit den Notenfolgen der gespeicherten Musikstücke vergleichen zu können, wenn die eingegebene Tonfolge bewußt oder unbewußt in eine andere Stimmlage transponiert wurde, ist bei einer zweckmäßigen Weiterbildung der Erfindung vorgesehen, daß zum Feststellen eines Übereinstimmungsfaktors $F_{i,l}$ zwischen einer eingegebenen und einer gespeicherten Tonfolge die Differenz der Tonhöhen h und Tondauern d der jeweiligen Notenfolgen miteinander verglichen werden.

[0022] Eine praktische Ausgestaltung der Erfindung zeichnet sich dadurch aus, daß beim Vergleich der Notenfolgen einer eingegebenen und einer gespeicherten Tonfolge die Notenfolge der eingegebenen Tonfolge nacheinander mit entsprechenden Teil-Notenfolgen der gespeicherten Tonfolgen verglichen wird, um jeweils einen Übereinstimmungsfaktor $f_i(x)$ zu ermitteln und daß der den höchsten Übereinstimmungsgrad anzeigenden Übereinstimmungsfaktor $F_{i,l} = f_i(x_l)$ der gespeicherten Tonfolge als Übereinstimmungsfaktor zugeordnet wird.

[0023] Für die Durchführung der Erfindung mittels Datenverarbeitungsanlagen ist es besonders zweckmäßig, wenn als Übereinstimmungsfaktor der kleinste Wert einer Funktion $f_i(x)$ ermittelt wird, die durch die folgende Gleichung gegeben ist:

$$f_i(x) = \alpha \sum_{l=0}^{N-1} |h(l) - m_h - (h_i(x+l) - m_{h_i}(x))| + \beta \sum_{l=0}^{N-1} |d(l) - m_d - (d_i(x+l) - m_{d_i}(x))|$$

wobei α und β Gewichtungsfaktoren sind, für die gilt: $0 < \alpha, \beta$ und $\alpha + \beta = 1$; $h(l)$ die Tonhöhe des l -ten Tons einer eingegebenen Tonfolge, m_h der Median der Tonhöhen der eingegebenen Tonfolge, $d(l)$ die Tondauer des l -ten Tons einer eingegebenen Tonfolge, m_d der Median der Tondauern der eingegebenen Tonfolge, $h_i(x)$ die Tonhöhe des x -ten Tons einer gespeicherten Tonfolge, $d_i(x)$ die Tondauer des x -ten Tons dieser eingespeicherten Tonfolge, $m_{h_i}(x)$ der Median der Tonhöhen im Intervall $h_i(x)$ bis $h_i(x + N - 1)$, $m_{d_i}(x)$ der Median der Tondauern im Intervall $d_i(x)$ bis $d_i(x + N - 1)$.

[0024] Um dem Benutzer die Auswahl des gesuchten Musikstücks weiter zu erleichtern, ist bei einer zweckmäßigen Weiterbildung der Erfindung vorgesehen, daß die auszugebenden Bezeichnungen der Tonfolgen entsprechend einem Übereinstimmungsgrad zwischen den zugehörigen gespeicherten Tonfolgen und der eingegebenen Tonfolge sortiert werden und daß die Ausgabe mit der Bezeichnung beginnt, deren Tonfolge der eingegebenen Tonfolge am ähnlichsten ist, wobei nur Bezeichnungen solcher Tonfolgen ausgegeben werden, deren Übereinstimmungsgrad größer als ein vorgegebener Wert ist.

[0025] Eine besonders vorteilhafte Ausgestaltung der Erfindung zeichnet sich dadurch aus, daß die Notenfolgen der Vielzahl von Tonfolgen zusammen mit entsprechenden Bezeichnungen der Tonfolgen in einer Datenbankdatei gespeichert sind, wobei zusammen mit den in der Datenbankdatei gespeicherten Notenfolgen kurze charakteristische Ausschnitte der jeweiligen Tonfolgen gespeichert sind.

[0026] Erfindungsgemäß wird also eine besondere Datenbankdatei bereitgestellt, in der die Notenfolgen der in einer Datenbank zur Verfügung stehenden Musikstücke zusammen mit entsprechenden Bezeichnungen, also mit den Titeln der Musikstücke, gespeichert sind, so daß bei dem Vergleich der Notenfolge der eingegebenen Tonfolge die Notenfolgen der Musikstücke nicht jedesmal neu erzeugt werden müssen, so daß die Suche nach dem gewünschten Musikstück wesentlich vereinfacht und beschleunigt werden kann. Neben der Bezeichnung des Musikstücks kann in dieser besonderen Datenbankdatei auch jeder Notenfolge noch ein kurzer charakteristischer Ausschnitt des jeweiligen Musikstücks zugeordnet sein, zum Beispiel im MIDI-Format, so daß auf die Datenbankdatei, in der Musikstücke als solche gespeichert sind, erst zugegriffen zu werden braucht, wenn der Benutzer sich für ein konkretes Musikstück entschieden hat.

[0027] Die Erfindung wird im folgenden beispielsweise unter Bezugnahme auf die Zeichnung näher erläutert. Es zeigen:

Figur 1 ein schematisches Blockschaltbild eines Kommunikationssystems zur Durchführung der erfindungsgemäßen Verfahren,

Figur 2 den zeitlichen Verlauf einer geglätteten Tonhöhenfrequenz und

Figur 3 den zeitlichen Verlauf einer entsprechend dem musikalischen Noten oder Tönen quantisierten Tonhöhenfrequenz.

[0028] Figur 1 zeigt beispielsweise ein Kommunikationssystem, bei dem ein Benutzer mit Hilfe eines Benutzerendgeräts, das beispielsweise als Mobiltelefon 10 ausgebildet ist, über eine Übertragungsstrecke 11 mit einer Datenbankstation 12 eines Diensteanbieters, die eine Musikstückdatenbank 13 umfaßt, kommunizieren kann, um Musikstücke, Videoclips und gegebenenfalls Filme oder dergleichen geliefert zu bekommen.

[0029] Das Mobiltelefon 10 besitzt in üblicher Weise ein Mikrofon 14 zur Sprach- und Toneingabe, dessen Ausgang über einen Analog/Digital-Wandler 15 an eine zentrale Verarbeitungsschaltung 16 angelegt ist. Die zentrale Verarbeitungsschaltung 16, die beispielsweise als Mikroprozessor ausgebildet sein kann, gibt an die Datenbankstation 12 des Diensteanbieters zu übertragende Daten an eine Sende- und Empfangseinheit 17 aus, an die eine Sende- und Empfangsantenne 18 angeschlossen ist, um Informationen über die Übertragungsstrecke 11 zu senden und von dieser zu empfangen.

[0030] Die Datenbankstation 12 des Diensteanbieters weist eine Sende- und Empfangseinheit 19 mit einer Sende- und Empfangsantenne 20 auf, um Daten von bzw. über die Übertragungsstrecke 11 empfangen bzw. senden zu können. Die Sende- und Empfangseinheit 19 ist mit einer zentralen Verarbeitungsschaltung 21 verbunden, die auf die Musikstückdatenbank 13 zugreifen kann, um ein angefordertes Musikstück zu dem Mobiltelefon 10 zu übertragen.

[0031] Für die Erkennung von Musikstücken ist eine Datenbankdatei 22 vorgesehen, in der zusammen mit den Bezeichnungen oder Titeln der einzelnen Musikstücke in der Musikstückdatenbank 13 den Musikstücken entsprechende Notenfolgen gespeichert sind. Zusammen mit den Bezeichnungen und Notenfolgen der Musikstücke können dabei noch charakteristische Ausschnitte aus den Musikstücken gespeichert sein.

[0032] Zur akustischen und optischen Ausgabe von Information weist das Mobiltelefon 10 einen Lautsprecher 23 und eine Anzeigevorrichtung 24 auf, die über entsprechende Treiberschaltungen 25 bzw. 26 mit der zentralen Verarbeitungsschaltung 16 verbunden sind.

[0033] Um ein bestimmtes Musikstück von einem Diensteanbieter anzufordern, gibt der Benutzer zunächst einen Ausschnitt des auszuwählenden bzw. gewünschten Musikstücks ein, in dem er einfach die ihm bekannte Melodie in das Mikrofon 14 singt. Die vom Mikrofon aufgezeichnete menschliche Stimme wird über den Analog/Digital-Wandler 15 digitalisiert und an die zentrale Verarbeitungsschaltung 16 geliefert, die somit den digitalisierten Frequenzverlauf der menschlichen Stimme empfängt.

[0034] Aus dem digitalisierten Frequenzverlauf der menschlichen Stimme wird mit Hilfe eines Tonhöhendetektors in der zentralen Verarbeitungsschaltung 16 der zeitliche Verlauf der Tonhöhenfrequenz der in das Mikrofon 14 gesun-

genen Tonfolge ermittelt. Als Tonhöhendetektor dient dabei beispielsweise der sogenannte SIFT (Simplified inverse filter tracking) Algorithmus, der besonders gut für relativ hohe Frauenstimmen geeignet ist, oder die sogenannte Cepstrum Tonhöhenabschätzung (Cepstrum pitch estimation) die sich für relativ tiefe männliche Stimmen eignet. Diese Verfahren sind dem zuständigen Fachmann vertraut und beispielsweise in dem Lehrbuch "Voice and Speech Processing", Thomas W. Parsons, New York, 1986, McGraw-Hill Book Company, erläutert.

[0035] Der ermittelte Verlauf der Tonhöhenfrequenz f_p wird dann mit Hilfe eines geeigneten Filters geglättet. Hierbei wird insbesondere ein Medianfilter verwendet, bei dem ein Filterfenster über die zu glättende Tonhöhenfrequenzkurve gleitet, um jeweils den Wert im Zentrum des Fensters durch den Median aller Werte im Fenster zu ersetzen. Eine derartige Medianfilterung ist ebenfalls bekannt und in dem obengenannten Lehrbuch erläutert.

[0036] Nach dem Glätten ergibt sich ein Verlauf der Tonhöhenfrequenz f_p , wie er in Figur 2 rein schematisch angedeutet ist. Somit liegt also ein geglätteter Verlauf der Tonhöhenfrequenzen der gesungenen Tonfolge über der Zeit vor, die im Idealfall mit dem Verlauf der Melodie im Frequenzbereich koinzidiert.

[0037] Da jedoch durch bewußtes und unbewußtes Transponieren der Melodie beim Singen durch den Benutzer sowie durch Rythmus- und Tempounterschiede Fehler oder Abweichungen der gesungenen Melodie von der gewünschten Melodie auftreten, wird der in Figur 2 gezeigte Verlauf der Tonhöhenfrequenzen entsprechend den Frequenzen der musikalischen Töne oder Noten quantisiert, so daß sich der in Figur 3 gezeigte quantisierte Verlauf der Tonhöhenfrequenzen f_p über der Zeit ergibt. Figur 3 zeigt dabei beispielsweise fünf verschiedene Töne mit unterschiedlichen Tondauern, denen jeweils ein bestimmter musikalischer Ton oder eine Note sowie eine bestimmte Tondauer zugeordnet werden können.

[0038] Nach der Quantisierung des Verlaufs der Tonhöhenfrequenz läßt sich die eingegebene, gesungene Tonfolge in eine bestimmte Anzahl N von Einzeltönen zerlegen. Jedem dieser Einzeltöne ist dabei ein musikalischer Ton entsprechend der musikalischen Skala zugeordnet. Ferner besitzt jeder der Einzeltöne eine bestimmte Tondauer, aus der sich eine entsprechende Notendauer ermitteln läßt.

[0039] Jeder Ton zeichnet sich also durch zwei Größen aus, nämlich durch die Tonhöhe oder Tonhöhenfrequenz, die durch den entsprechenden musikalischen Ton bzw. die entsprechende Note bezeichnet wird, und durch die Tondauer, die in noch zu beschreibender Weise entsprechend der musikalischen Notendauer quantisiert wird. Damit läßt sich jede Tonfolge aus N Tönen durch einen Tonhöhenvektor

$$\mathbf{h} = (h_1, h_2, \dots, h_N)^T$$

und durch einen Tondauervektor

$$\mathbf{d} = (d_1, d_2, \dots, d_N)^T$$

beschreiben. Die Werte h_i können dabei einfach ganze Zahlen sein, die entsprechend der folgenden Tabelle die jeweiligen musikalischen Töne oder Noten repräsentieren.

Note	A'	Ais'	H'	C'	Cis'	D'	Dis'	E'	F'	Fis'	G'	Gis'	A''	Ais''	H''
Zahl	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14

[0040] In entsprechender Weise kann jeder Notendauer $1/32$, $1/16$, $1/8$, $1/4$, $1/2$, 1 eine entsprechende Zahl zugeordnet werden, wobei zweckmäßiger Weise für die kürzeste Note die Dauer 1 gesetzt wird. Eine $1/4$ -Note erhält dann die Dauer 8, eine $1/2$ -Note die Dauer 16 und die ganze Note die Dauer 32. Um den einzelnen Tondauern eine musikalische Notendauer zuordnen zu können, wird der Median der Tondauerverteilung ermittelt und gleich einer $1/4$ -Note gesetzt. Ausgehend von dem Median werden dann Zeitintervalle festgelegt, die den einzelnen Notendauern entsprechen.

[0041] Die gesungene Tonfolge liegt nunmehr als Notenfolge vor, die sich durch zwei äußerst einfache Vektoren beschreiben lassen.

[0042] Die Umsetzung der Tonfolge in die die Notenfolge beschreibenden Vektoren kann dabei in der zentralen Verarbeitungsschaltung 21 der Datenbankstation 12 des Diensteanbieters durchgeführt werden. Um jedoch die Übertragungsstrecke 11 möglichst wenig zu belasten, um also die entsprechenden Sendekanäle möglichst wenig zu blockieren, wird diese Umsetzung bereits im Mobiltelefon 10 von der zentralen Verarbeitungsschaltung 16 vorgenommen, so daß nur der Tonhöhenvektor und der Notendauervektor zur Datenbankstation 12 des Diensteanbieters übertragen

zu werden brauchen.

[0043] In der Datenbankstation 12 sind in der Datenbankdatei 22 die Musikstücke als Notenfolgen gespeichert, die ebenfalls durch einen entsprechenden Tonhöhenvektor $\mathbf{h}_i = (h_{i1}, h_{i2}, \dots, h_{ix}, \dots, h_{iM})$ und Tondauervektoren $\mathbf{d}_i = (d_{i1}, d_{i2}, \dots, d_{ix}, \dots, d_{iM})$ beschrieben werden. Der Index i bezeichnet dabei das jeweilige Musikstück und M die Anzahl der Töne oder Noten.

[0044] Um auch bewußt oder unbewußt transponierte eingegebene Tonfolgen mit den Musikstücken vergleichen zu können, werden nicht die jeweiligen Notenfolgen unmittelbar miteinander verglichen, sondern nur der Relativverlauf innerhalb der beiden Notenfolgen. Hierzu werden jeweils die Differenzen zwischen den einzelnen Tonhöhen miteinander verglichen. Es wird also für jede Notenfolge der Median festgestellt, um den Abstand der einzelnen Töne vom Median zu ermitteln und mit dem Abstand des entsprechenden anderen Tons der anderen Notenfolge von deren Median zu vergleichen. Da die Notenfolge des Musikstücks typischerweise wesentlich länger ist, als die Notenfolge, die beispielsweise durch Singen eingegeben wurde, wird für die Notenfolge des Musikstücks jeweils der Median eines entsprechenden Teilabschnitts dieser Notenfolge verwendet.

[0045] Beim praktischen Vergleich der Notenfolge einer eingegebenen Tonfolge mit den Notenfolgen der Musikstücke wird eine Funktion $f_i(x)$ berechnet, deren Verlauf anzeigt, wie die Notenfolge der eingegebenen Tonfolge mit den einzelnen Abschnitten übereinstimmt. Diese Abweichungsfunktion errechnet sich nach der folgenden Gleichung:

$$f_i(x) = \alpha \sum_{l=0}^{N-1} |h(l) - m_h - (h_i(x+l) - m_{h_i}(x))| + \beta \sum_{l=0}^{N-1} |d(l) - m_d - (d_i(x+l) - m_{d_i}(x))|$$

[0046] Hierbei sind α und β Gewichtungsfaktoren, die den Einfluß der Melodie bzw. des Rhythmus auf den Übereinstimmungsfaktor beschreiben. Für α und β gilt dabei: $0 < \alpha, \beta$; $\alpha + \beta = 1$. $h_i(x)$ und $d_i(x)$ bezeichnen die Tonhöhe und die Tondauer des x -ten Tons im Vektor $\mathbf{h}_i$ bzw. $\mathbf{d}_i$. $m_{h_i}(x)$ und $m_{d_i}(x)$ bezeichnen jeweils den Median der Tonhöhen bzw. Tondauern im Intervall von $h_i(x)$ bis $h_i(x + N - 1)$ bzw. $d_i(x)$ bis $d_i(x + N - 1)$. $h(l)$ und $d(l)$ bezeichnen die Tonhöhe bzw. Tondauer des l -ten Tons im Vektor $\mathbf{h}$ bzw. $\mathbf{d}$. In entsprechender Weise bezeichnen m_h und m_d den Median der Tonhöhen bzw. Tondauern im Vektor $\mathbf{h}$ bzw. im Vektor $\mathbf{d}$.

[0047] Sowohl für die Tonhöhen als auch für die Tondauern wird jeweils die Summe der Differenzen zwischen den jeweiligen Abständen zum entsprechenden Median berechnet, die im Idealfall, wenn also die Notenfolgen exakt übereinander passen, gleich 0 wird.

[0048] Nachdem die Funktion $f_i(x)$ für alle Werte x berechnet wurde, wenn also die Notenfolge der eingegebenen Tonfolge mit sämtlichen möglichen Abschnitten der Notenfolge eines Musikstücks in der durch die obige Gleichung beschriebene Weise verglichen wurde, wird der kleinste Wert der Funktion $f_i(x)$ festgestellt. Der zugehörige Wert x_i beschreibt somit den Abschnitt der Notenfolge der den vom Benutzer gesungenen Abschnitt des Musikstücks (möglicherweise) entspricht. Der zugehörige Wert der Funktion $f_i(x)$ wird dann als Abweichungsfaktor $F_{i,i} = f_i(x_i)$ gespeichert.

[0049] Sobald die Notenfolge der eingegebenen Tonfolge mit sämtlichen Notenfolgen der einzelnen Musikstücke verglichen wurde, werden die Bezeichnungen oder Titel der Musikstücke entsprechend den ermittelten Übereinstimmungsfaktoren $F_{i,i}$ sortiert, wobei mit dem die höchste Übereinstimmung bezeichnenden kleinsten Abweichungsfaktor begonnen wird.

[0050] Um anschließend dem Benutzer die Musikstücke in der ermittelten Reihenfolge zu präsentieren, werden diese von der Datenbankstation 12 zum Mobiltelefon 10 übertragen, wo die Titel auf der Anzeigevorrichtung 24 angezeigt werden während charakteristische Ausschnitte der Musikstücke über den Lautsprecher 23 ausgegeben werden können. Hierbei wird zweckmäßigerweise die Anzahl der übertragenen Bezeichnungen begrenzt. Die Begrenzung kann dabei im einfachsten Falle dadurch erfolgen, daß nur, je nach Anzeige- und Speicherkapazität dem Mobiltelefon eine begrenzte feste Anzahl von Titeln der Musikstücke übertragen wird. Es ist jedoch auch möglich, die Begrenzung im Hinblick auf den Abweichungsfaktor vorzunehmen, so daß nur Titel von Musikstücken zum Mobiltelefon übertragen und dem Benutzer angezeigt werden, deren Abweichungsfaktor einen vorbestimmten Schwellwert nicht übersteigt. Ein derartiger Schwellwert kann allgemein festgelegt oder entsprechend der Abweichungsfaktorverteilung ermittelt werden.

[0051] Die vorliegende Erfindung ermöglicht somit die Erkennung von Musikstücken in der Datenbankstation eines Diensteanbieters, wobei ein Benutzer nur einen Teil eines gewünschten Musikstücks singt, wenn ihm der Titel dieses Liedes oder Musikstücks nicht bekannt ist. Nach der Erkennung des Musikstücks, oder einer Reihe von möglichen Musikstücken wird oder werden der bzw. die Titel gegebenenfalls zusammen mit charakteristischen Ausschnitten der Musikstücke zum Benutzer übertragen, so daß dieser daraus das gewünschte Musikstück auswählen kann. Nach der Auswahl wird dann das komplette Musikstück über elektronische Kommunikationswege (Internet, zellulARES Mobilfunknetz, wie im beschriebenen Ausführungsbeispiel, oder ähnliches) zugesandt und der Benutzer kann sich das Musik-

stück dauerhaft auf einem geeigneten Speichermedium (CD, Speicherbaustein, Magnetband usw.) abspeichern und abspielen.

[0052] Für den Vergleich der eingegebenen Tonfolge, also eines gesungenen Abschnitts des gewünschten Musikstücks mit den Musikstücken in der Datenbankstation des Diensteanbieters ist in der Datenbankstation 12 eine gesonderte Datenbankdatei 22 vorgesehen, in der die Titel bzw. Bezeichnungen der einzelnen Musikstücke mit den zugeordneten Notenfolgen gespeichert sind, so daß das Auffinden der gewünschten Musikstücke wesentlich erleichtert und die Erkennung beschleunigt ist.

Patentansprüche

1. Verfahren zum Erkennen einer Tonfolge insbesondere eines Musikstücks, bei dem

- die Töne der zu erkennenden Tonfolge in eine Notenfolge umgesetzt werden,
- zum Suchen der zu erkennenden Tonfolge deren Notenfolge nacheinander mit entsprechenden Notenfolgen einer Vielzahl von Tonfolgen verglichen wird, und
- Bezeichnungen der Tonfolge oder -folgen ausgegeben werden, deren Notenfolge bzw. -folgen mit der Notenfolge der zu erkennenden Tonfolge in vorbestimmter Weise übereinstimmt bzw. übereinstimmen.

2. Verfahren zum Auswählen einer Tonfolge, insbesondere eines Musikstücks, bei dem

- eine zumindest einem Abschnitt der auszuwählenden Tonfolge zumindest teilweise entsprechende Tonfolge eingegeben wird,
- die Töne der eingegebenen Tonfolge in eine Notenfolge umgesetzt werden,
- zum Suchen der auszuwählenden Tonfolge deren Notenfolge nacheinander mit entsprechenden Notenfolgen einer Vielzahl von Tonfolgen verglichen wird, um Bezeichnungen einer oder mehrerer Tonfolgen zu ermitteln, deren Notenfolge oder -folgen mit der Notenfolge der auszuwählenden Tonfolge in vorbestimmter Weise übereinstimmt bzw. übereinstimmen, und
- die ermittelten Bezeichnungen als Liste ausgegeben werden, so daß ein Benutzer anhand der Bezeichnungsliste die gewünschte Tonfolge auswählen kann.

3. Verfahren nach Anspruch 2, dadurch gekennzeichnet, daß die in einem Benutzerendgerät (10) eingegebene, der auszuwählenden Tonfolge entsprechende Tonfolge zu einer Datenbankstation (12) übertragen wird, in der die Liste der Bezeichnungen einer oder mehrerer der auszuwählenden Tonfolge ähnlichen Tonfolgen ermittelt wird, und

- die Bezeichnungsliste zur Ausgabe an das Benutzerendgerät (10) übertragen wird.

4. Verfahren nach Anspruch 2, dadurch gekennzeichnet, daß die in ein Benutzerendgerät (10) eingegebene, der auszuwählenden Tonfolge entsprechende Tonfolge im Benutzerendgerät (10) in eine Notenfolge umgesetzt wird, die Notenfolge zu einer Datenbankstation (12) übertragen wird, in der die Liste der Bezeichnungen einer oder mehrerer der auszuwählenden Tonfolge ähnlichen Tonfolgen ermittelt wird, und

- die Bezeichnungsliste zur Ausgabe an das Benutzerendgerät (10) übertragen wird.

5. Verfahren nach Anspruch 3 oder 4, dadurch gekennzeichnet, daß zusammen mit jeder Bezeichnung ein für die jeweilige Tonfolge charakteristischer kurzer Ausschnitt der Tonfolge zur Ausgabe an das Benutzerendgerät (10) übertragen wird.

6. Verfahren nach Anspruch 2 bis 5, dadurch gekennzeichnet, daß die Tonfolge zur Eingabe in das Benutzerendgerät (10) vom Benutzer gesungen wird.

7. Verfahren nach einem der vorstehenden Ansprüche, dadurch gekennzeichnet, daß zum Umsetzen einer Tonfolge in eine Notenfolge

- für jeden Ton der Tonfolge die Tonhöhenfrequenz f_p und die Tondauer ermittelt wird, und
- jedem Ton eine musikalische Note entsprechend seiner Tonhöhenfrequenz f_p und eine musikalisch quantisierte Notendauer entsprechend einer Tondauerverteilung der Tonfolge zugeordnet wird.

8. Verfahren nach Anspruch 7, dadurch gekennzeichnet, daß zur Festlegung der Notendauer der Töne zunächst der

Median der Tondauerverteilung ermittelt wird, und die Tondauer des Medians mit der Notendauer einer 1/4-Note gleichgesetzt wird, sowie jedem Ton durch Vergleich seiner Tondauer mit der ermittelten Notendauer einer 1/4-Note eine entsprechende musikalische quantisierte Notendauer (1/32, 1/16, 1/8, 1/4, 1/2, 1) zugeordnet wird.

- 5 9. Verfahren nach Anspruch 7 oder 8, dadurch gekennzeichnet, daß jede Tonfolge durch einen Tonhöhenvektor **h**, der sich aus den einzelnen Noten oder musikalischen Tönen zusammensetzt, und einen Tondauervektor **d** dargestellt wird, der sich aus den musikalischen quantisierten Notendauern der einzelnen Töne zusammensetzt.
- 10 10. Verfahren nach einem der vorstehenden Ansprüche, dadurch gekennzeichnet, daß zum Feststellen eines Abweichungsfaktors $F_{i,1}$ zwischen einer eingegebenen und einer gespeicherten Tonfolge die Differenz der Tonhöhen und Tondauern der jeweiligen Notenfolgen miteinander verglichen werden.
- 15 11. Verfahren nach einem der vorstehenden Ansprüche, dadurch gekennzeichnet, daß beim Vergleich der Notenfolgen einer eingegebenen und einer gespeicherten Tonfolge die Notenfolge der eingegebenen Tonfolge nacheinander mit entsprechenden Teil-Notenfolgen der gespeicherten Tonfolgen verglichen wird, um jeweils einen Abweichungsfaktor $f_i(x)$ zu ermitteln und daß der den höchsten Übereinstimmungsgrad anzeigende kleinste Abweichungsfaktor $F_{i,1} = f_i(x_i)$ der gespeicherten Tonfolge als Abweichungsfaktor zugeordnet wird.
- 20 12. Verfahren nach Anspruch 10 und 11, dadurch gekennzeichnet, daß als Abweichungsfaktor der kleinste Wert einer Funktion $f_i(x)$ ermittelt wird, die durch die folgende Gleichung gegeben ist:

$$f_i(x) = \alpha \sum_{l=0}^{N-1} |h(l) - m_h - (h_i(x+l) - m_{h_i}(x))| + \beta \sum_{l=0}^{N-1} |d(l) - m_d - (d_i(x+l) - m_{d_i}(x))|$$

25 wobei α und β Gewichtungsfaktoren sind, für die gilt: $0 < \alpha, \beta$ und $\alpha + \beta = 1$; $h(l)$ die Tonhöhe des l-ten Tons einer eingegebenen Tonfolge, m_h der Median der Tonhöhen der eingegebenen Tonfolge, $d(l)$ die Tondauer des l-ten Tons einer eingegebenen Tonfolge, m_d der Median der Tondauern der eingegebenen Tonfolge, $h_i(x)$ die Tonhöhe des x-ten Tons einer gespeicherten Tonfolge, $d_i(x)$ die Tondauer des x-ten Tons dieser eingespeicherten Tonfolge, $m_{h_i}(x)$ der Median der Tonhöhen im Intervall $h_i(x)$ bis $h_i(x + N - 1)$, $m_{d_i}(x)$ der Median der Tondauern im Intervall $d_i(x)$ bis $d_i(x + N - 1)$.

- 30 13. Verfahren nach einem der vorstehenden Ansprüche, dadurch gekennzeichnet, daß die auszugebenden Bezeichnungen der Tonfolgen entsprechend einem Übereinstimmungsgrad zwischen den zugehörigen gespeicherten Tonfolgen und der eingegebenen Tonfolge sortiert werden und daß die Ausgabe mit der Bezeichnung beginnt, deren Tonfolge der eingegebenen Tonfolge am ähnlichsten ist.
- 35 14. Verfahren nach einem der vorstehenden Ansprüche, dadurch gekennzeichnet, daß nur Bezeichnungen solcher Tonfolgen ausgegeben werden, deren Übereinstimmungsgrad größer als sein vorgegebener Wert ist.
- 40 15. Verfahren nach einem der vorstehenden Ansprüche, dadurch gekennzeichnet, daß die Notenfolgen der Vielzahl von Tonfolgen zusammen mit entsprechenden Bezeichnungen der Tonfolgen in einer Datenbankdatei (22) gespeichert sind.
- 45 16. Verfahren nach Anspruch 15, dadurch gekennzeichnet, daß zusammen mit den in der Datenbankdatei (22) gespeicherten Notenfolgen kurze charakteristische Ausschnitte der jeweiligen Tonfolgen gespeichert sind.

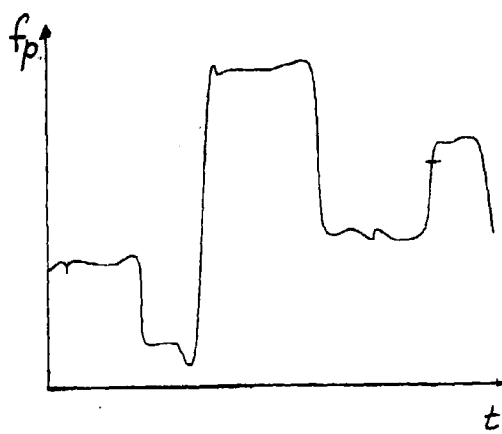
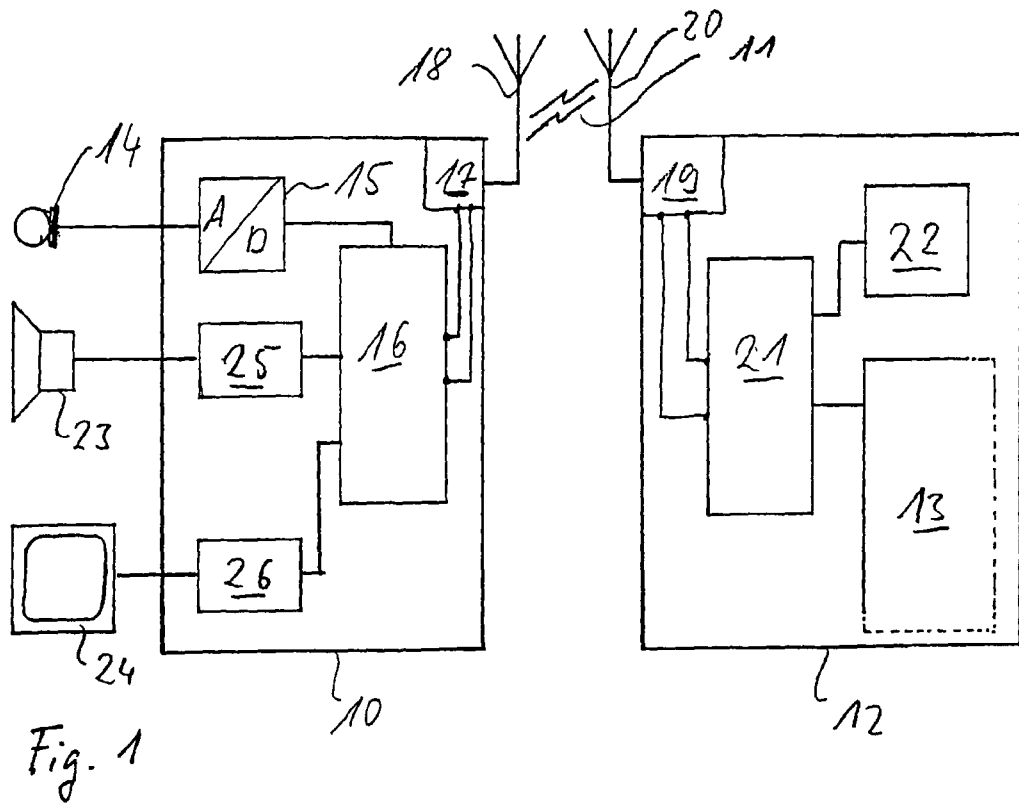


Fig. 2

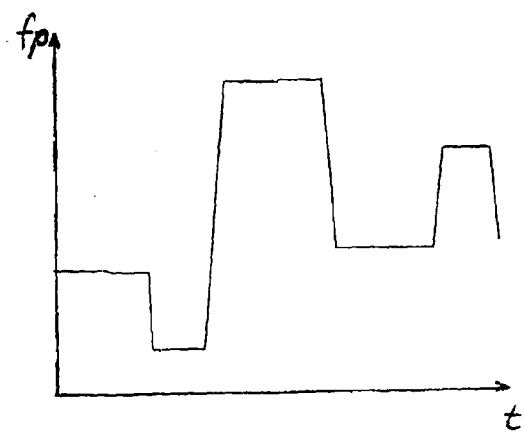


Fig. 3



Europäisches
Patentamt

EUROPÄISCHER RECHERCHENBERICHT

Nummer der Anmeldung
EP 00 12 0872

EINSCHLÄGIGE DOKUMENTE			
Kategorie	Kennzeichnung des Dokuments mit Angabe, soweit erforderlich, der maßgeblichen Teile	Betrifft Anspruch	KLASSIFIKATION DER ANMELDUNG (Int.Cl.7)
X	US 5 874 686 A (GHIA S ASIF U ET AL) 23. Februar 1999 (1999-02-23) * Spalte 2, Zeile 27 - Zeile 59 * * Spalte 6, Zeile 60 - Spalte 7, Zeile 8; Abbildung 1 *	1-4, 6, 13-15	G10H1/00
Y	---	5, 16	
Y	US 5 739 451 A (EGYUD JULES ET AL) 14. April 1998 (1998-04-14) * Spalte 3, Zeile 62 - Spalte 4, Zeile 26 * * Spalte 5, Zeile 40 - Zeile 61 * * Spalte 7, Zeile 5 - Zeile 27; Abbildung 1 *	5, 16	
A	---	1-4	
A	DE 196 52 225 A (RIECK HARALD) 25. Juni 1998 (1998-06-25) * Spalte 1, Zeile 61 - Spalte 2, Zeile 34 *	1-4, 15	
A	EP 0 944 033 A (SONODA TOMONARI) 22. September 1999 (1999-09-22) * Absatz '0015! - Absatz '0028!; Abbildungen 3, 4 *	1-16	RECHERCHIERTE SACHGEBIETE (Int.Cl.7) G10H
A	US 5 963 957 A (HOFFBERG MARK B) 5. Oktober 1999 (1999-10-05) * Spalte 1, Zeile 66 - Spalte 2, Zeile 51; Abbildung 1 *	1-4	
Der vorliegende Recherchenbericht wurde für alle Patentansprüche erstellt			
Recherchenort DEN HAAG		Abschlußdatum der Recherche 15. Februar 2001	Prüfer Pulluard, R
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	

EPO FORM 1503.03 B2 (P04C03)

**ANHANG ZUM EUROPÄISCHEN RECHERCHENBERICHT
 ÜBER DIE EUROPÄISCHE PATENTANMELDUNG NR.**

EP 00 12 0872

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

Die Angaben über die Familienmitglieder entsprechen dem Stand der Datei des Europäischen Patentamts am

Diese Angaben dienen nur zur Unterrichtung und erfolgen ohne Gewähr.

15-02-2001

Im Recherchenbericht angeführtes Patentedokument	Datum der Veröffentlichung	Mitglied(er) der Patentfamilie	Datum der Veröffentlichung
US 5874686 A	23-02-1999	KEINE	
US 5739451 A	14-04-1998	KEINE	
DE 19652225 A	25-06-1998	KEINE	
EP 0944033 A	22-09-1999	JP 11272274 A	08-10-1999
		JP 2000187671 A	04-07-2000
		US 6121530 A	19-09-2000
US 5963957 A	05-10-1999	EP 0919033 A	02-06-1999
		WO 9849630 A	05-11-1998
		JP 2000513846 T	17-10-2000

EPO FORM P0481

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