



BUNDESPATENTGERICHT

IM NAMEN DES VOLKES

2 Ni 68/20 (EP)

(AktENZEICHEN)

URTEIL

Verkündet am
5. Mai 2022

...

In der Patentnichtigkeitsache

...

...

betreffend das europäische Patent 1 290 889
(DE 601 17 376)

hat der 2. Senat (Nichtigkeitssenat) des Bundespatentgerichts aufgrund der mündlichen Verhandlung vom 5. Mai 2022 durch die Vorsitzende Richterin Hartlieb sowie die Richter Dipl.-Phys. Dr. Forkel, Dipl.-Ing. Hoffmann, Dr. Himmelmann und Dipl.-Phys. Dr. Städele für Recht erkannt:

- I. Das europäische Patent EP 1 290 889 wird mit Wirkung für das Hoheitsgebiet der Bundesrepublik Deutschland im Umfang der Ansprüche 1, 2, 13 und 14 für nichtig erklärt.
- II. Die Kosten des Rechtsstreits trägt die Beklagte.
- III. Das Urteil ist gegen Sicherheitsleistung in Höhe von 120 % des zu vollstreckenden Betrages vorläufig vollstreckbar.

Tatbestand

Die Beklagte ist Inhaberin des auch mit Wirkung für die Bundesrepublik Deutschland erteilten europäischen Patents 1 290 889 (Streitpatent), das am 22. Februar 2006 in englischer Sprache veröffentlicht wurde. Das Streitpatent hat den Anmeldetag 7. Mai 2001 und nimmt die Priorität US 60/210,440 P vom 8. Juni 2000 und die Priorität US 09/785,375 vom 16. Februar 2001 in Anspruch. Die Anmeldeunterlagen des Streitpatents wurden am 12. März 2003 mit der EP 1 290 889 A1 offengelegt.

Das Streitpatent trägt die Bezeichnung „SYSTEM AND METHOD OF VOICE RECOGNITION NEAR A WIRELINE NODE OF A NETWORK SUPPORTING CABLE TELEVISION AND/OR VIDEO DELIVERY = SYSTEM UND VERFAHREN ZUR SPRACHERKENNUNG BEI EINEM DRAHTGEBUNDENEN KNOTEN EINES NETZES FÜR KABELFERNSEHEN UND/ODER VIDEOVERTEILUNG“. Die dem Streitpatent zugrundeliegende internationale Anmeldung wurde am 13. Dezember 2001 unter der Publikationsnummer WO 01 / 95625 A1 veröffentlicht. Das am 7. Mai 2021 durch Zeitablauf erloschene Patent wird beim Deutschen Patent- und Markenamt unter dem Aktenzeichen 601 17 376.7 geführt und umfasst 31 Ansprüche. Die Patentansprüche 2 bis 12 und 14 bis 31 sind abhängige Patentansprüche des Streitpatents.

Die Klägerin begehrt die Nichtigkeitsklärung des deutschen Teils des Streitpatents im Umfang der Ansprüche 1, 2, 13 und 14. Die Beklagte verteidigt das Streitpatent in vollem Umfang und hilfsweise beschränkt mit acht Hilfsanträgen.

Die Klägerin stützt ihre Klage auf die Nichtigkeitsgründe der mangelnden Patentfähigkeit mit Blick auf fehlende Neuheit und fehlende erfinderische Tätigkeit sowie der unzureichenden Offenbarung und der unzulässigen Erweiterung.

Zur Stützung ihres Vorbringens hat die Klägerin die folgenden Dokumente genannt:

- NK1 EP 1 290 889 B1 (Streitpatentschrift);
- NK1a DE 601 17 376 T2 (dt. Übersetzung der Streitpatentschrift);
- NK2 ursprünglich eingereichte Anmeldeunterlagen;
- NK3 US-Prioritätsanmeldung 60/210,440;
- NK4 US-Prioritätsanmeldung 09/785,375;
- NK5 EPA-Patentregisterauszug mit Datumsangabe „25.10.2020“;
- NK6 Klageschrift im Verletzungsverfahren;
- NK6a Replik der Verletzungsklägerin im Verletzungsverfahren;
- NK7 Merkmalsgliederung des Patentanspruchs 1 des Streitpatents;
- NK8 US 5 446 489 A;
- NK9 US 5 774 859 A;

- NK10 EP 0 782 337 A2;
- NK11 WO 00 / 24198 A1;
- NK12 „Speech Recognition Methods for Controlling Cable Television“, IBM Technical Disclosure Bulletin, Vol. 38, Nr. 8, Seiten 285 – 287, August 1995;
- NK13 „BBN’s Voice Navigation for Time-Warner’s FSN“, Telemedia News & Views 2.12 (Dezember 1994);
- NK13a vollständige Ausgabe der „Telemedia News & Views“ vom Dezember 1994, 9 Seiten;
- NK14 Price Colman, „The Power of Speech“, Auszug aus „Convergence™ - The Magazine for the new communications industry“, Seiten 16 - 23, August 1995;
- NK15 WO 00 / 58 942 A2;
- NK16 Digital Video Broadcasting (DVB) - Interaction channel for Cable TV distribution systems (CATV). European Telecommunications Standards Institute, Sophia Antipolis, France, Januar 1998;
- NK17 EP 0 872 827 A2;
- NK18 WO 97 / 49 242 A1;
- NK18a Stellungnahme der Beklagten im GB-Verletzungsverfahren;
- NK18b geänderte Stellungnahme der Beklagten im GB-Verletzungsverfahren;
- NK18c geänderter Claim 13 aus dem GB-Verletzungsverfahren;
- NK18d Schreiben der Prozessbevollmächtigten der Beklagten in GB;
- NK18e Entscheidung des High Court im GB-Verletzungsverfahren;
- NK19 Data-Over-Cable Service Interface Specifications - Radio Frequency Interface Specification. Cable Television Laboratories, Inc., mit Hinweis „Copyright 1999, 2000“ und Datumsangabe „April 7, 2000“, 272 Seiten (eingereicht mit Schriftsatz vom 28. August 2021);
- NK19 WO 97 / 13 368 A1 (eingereicht mit Schriftsatz vom 4. April 2022);
- NK20 US 5 758 259 A;
- NK21 US 5 861 906 A;
- NK22 WO 00 / 44 173 A1.

Die Klägerin behauptet, die Gegenstände der Ansprüche der Hilfsanträge 1 bis 8 seien nicht patentfähig.

Die Klägerin stellt den Antrag,

das europäische Patent EP 1 290 889 im Umfang der Ansprüche 1, 2, 13 und 14 mit Wirkung für das Hoheitsgebiet der Bundesrepublik Deutschland für nichtig zu erklären.

Die Beklagte stellt den Antrag,

die Klage abzuweisen,

hilfsweise das europäische Patent EP 1 290 889 unter Klageabweisung im Übrigen dadurch teilweise für nichtig zu erklären, dass seine angegriffenen Patentansprüche die Fassung eines der Hilfsanträge 1 bis 7 vom 17. Februar 2022 und Hilfsantrag 8 vom 5. Mai 2022 in dieser Reihenfolge, erhalten.

Die Beklagte erklärt, dass sie die Patentansprüche gemäß Hauptantrag und Hilfsanträgen als jeweils geschlossene Anspruchssätze ansieht, die jeweils insgesamt beansprucht werden.

Die Beklagte tritt der Argumentation der Klägerin in allen wesentlichen Punkten entgegen und vertritt die Auffassung, dass die angegriffenen Ansprüche neu seien und auf einer erfinderischen Tätigkeit beruhen würden, hinreichend offenbart und nicht unzulässig erweitert seien. Das Streitpatent sei jedenfalls in der Fassung eines der Hilfsanträge patentfähig.

Mit Schriftsatz vom 3. Mai 2022 hat die Beklagte auf die Stellungnahme der Klägerin vom 27. April 2022 erwidert und die darin enthaltenen neuen Angriffe auf Basis von (angeblich) noch nicht verwendeten Druckschriften als verspätet gerügt. Der umfassendste neue Angriff basiere auf der Druckschrift NK 22.

Der erteilte Patentanspruch 1 lautet in der Verfahrenssprache Englisch gemäß EP 1 290 889 B1 (mit an die Anlage NK7 der Klägerin angelehnter Merkmalsgliederung):

- 1.1 A method of using a back channel
 - 1.1.1 containing a multiplicity of identified speech channels from a multiplicity of user sites (1100)
 - 1.1.2 presented to a speech recognition system (3200) at a wireline node (1300) of a network supporting at least one of cable television delivery and video delivery,
- 1.2 comprising the steps of:
 - 1.2.1 receiving said back channel to create a received back channel,
 - 1.2.2 partitioning said received back channel into a multiplicity of received identified speech channels;
 - 1.2.3 processing each of said multiplicity of said received identified speech channels with said speech recognition system to create corresponding recognized speech content for each received identified speech channel;
 - 1.2.4 responding to said recognized speech content to create a recognized speech content response that is unique for each of said received identified speech channels; and
 - 1.2.5 individually controlling the delivery of entertainment and information services to each user site (1100) in accordance with said recognized speech.

Der erteilte Patentanspruch 1 lautet in der Verfahrenssprache Deutsch gemäß EP 1 290 889 B1 (mit an die Anlage NK7 der Klägerin angelehnter Merkmalsgliederung):

- 1.1 Verfahren zur Verwendung eines Rückkanals,
 - 1.1.1 der eine Vielzahl von erkannten Sprachkanälen von einer Vielzahl von Teilnehmerstandorten (1100) enthält,
 - 1.1.2 die in ein Spracherkennungssystem (3200) an einem Leitungs-Knoten (1300) eines Netzwerks, das mindestens einen der Dienste Kabelfernsehen oder Video-Verteildienst unterstützt, eingegeben werden
- 1.2 und das folgende Schritte umfasst:

- 1.2.1 Empfang des Rückkanals zur Erzeugung eines empfangenen Rückkanals;
- 1.2.2 Aufteilung des empfangenen Rückkanals in eine Vielzahl empfangener erkannter Sprachkanäle;
- 1.2.3 Verarbeitung jedes aus der Vielzahl empfangener erkannter Sprachkanäle mit dem Spracherkennungssystem, um für jeden empfangenen erkannten Sprachkanal einen entsprechenden erkannten Sprachinhalt zu erzeugen;
- 1.2.4 Reaktion auf den erkannten Sprachinhalt, um eine Reaktion auf den erkannten Sprachinhalt zu erzeugen, die für jeden der erkannten Sprachkanäle eindeutig ist; und
- 1.2.5 Individuelle Kontrolle der Lieferung von Unterhaltungs- und Informationsdiensten zu jedem Teilnehmerstandort (1100) in Übereinstimmung mit der erkannten Sprache.

Der erteilte Patentanspruch 13 lautet in der Verfahrenssprache Englisch gemäß EP 1 290 889 B1 (mit einer Merkmalsgliederung):

- 13.1 A system supporting speech recognition for a network that supports at least one of cable television delivery (1100) and video delivery (1100) to a multiplicity of users,
said system comprising:
- 13.2 a speech recognition system (3200) coupled to a wireline node (1300) being part of said network for receiving a back channel from a multiplicity of user sites (1100) coupled to said network;
- 13.3 a back channel receiver (1322) for receiving said back channel to create a received back channel;
- 13.4 a speech channel partitioner (1580; 2012) for partitioning said received back channel into a multiplicity of received identified speech channels;
and
- 13.5 a processor (1520) for executing a program residing in a memory accessibly coupled to said processor (1510);
wherein said processor (1520) comprises

- 13.5.1 means for processing each of said multiplicity of said received identified speech channels within said speech recognition system (3200) to create corresponding recognized speech content for each received identified speech channel;
- 13.5.2 means for responding to said recognized speech content to create a recognized speech content response that is unique for each of said multiplicity of received identified speech channels; and
- 13.5.3 means for individually controlling the delivery of entertainment and information services to each user site (1100) in accordance with said recognized speech.

Der erteilte Patentanspruch 13 lautet in der Verfahrenssprache Deutsch gemäß EP 1 290 889 B1 (mit einer Merkmalsgliederung):

- 13.1 System, das die Spracherkennung für ein Netzwerk unterstützt, welches mindestens einen der Dienste zur Lieferung von Kabelfernsehen (1100) oder einen Video-Verteildienst (1100) für eine Vielzahl von Teilnehmern unterstützt,
wobei das System folgendes enthält:
- 13.2 Ein Spracherkennungssystem (3200), das mit einem Leitungs-Knoten (1300) gekoppelt ist, der Teil des Netzwerks ist, um einen Rückkanal von einer Vielzahl von Teilnehmerstandorten (1100), die mit dem Netzwerk gekoppelt sind, zu empfangen;
- 13.3 Einen Rückkanal-Empfänger (1322) zum Empfang des Rückkanals, um einen empfangenen Rückkanal zu erzeugen;
- 13.4 Einen Sprachkanal-Aufteiler (1580; 2012) zur Aufteilung des empfangenen Rückkanals in eine Vielzahl von empfangenen erkannten Sprachkanälen; und
- 13.5 Einen Prozessor (1520) zur Ausführung eines Programms, das sich in einem Speicher befindet, der zugänglich mit dem Prozessor (1510) gekoppelt ist;
wobei der Prozessor (1520) folgendes enthält:

- 13.5.1 Mittel zur Verarbeitung jedes aus der Vielzahl empfangener erkannter Sprachkanäle in dem Spracherkennungssystem (3200), um für jeden empfangenen erkannten Sprachkanal einen entsprechenden erkannten Sprachinhalt zu erzeugen;
- 13.5.2 Mittel zur Reaktion auf den erkannten Sprachinhalt, um eine Reaktion auf den erkannten Sprachinhalt zu erzeugen, die für jeden der erkannten Sprachkanäle eindeutig ist; und
- 13.5.3 Mittel zur individuellen Kontrolle der Lieferung von Unterhaltungs- und Informationsdiensten zu jedem Teilnehmerstandort (1100) in Übereinstimmung mit der erkannten Sprache.

Der erteilte Patentanspruch 2 lautet in der Verfahrenssprache Englisch gemäß EP 1 290 889 B1:

- 2. The method of Claim 1.
further comprising at least one of the steps of:
determining said associated user site (1100) from said received identified speech channel;
determining said associated user site (1100) from said recognized speech content;
determining said associated user site (1100) from said recognized speech content and a speaker identification library;
determining said associated user site (1100) from said recognized speech content and a speech recognition library; and
determining said associated user site (1100) from an identification within said speech channel.

Der erteilte Patentanspruch 2 lautet in der Verfahrenssprache Deutsch gemäß EP 1 290 889 B1:

- 2. Verfahren nach Anspruch 1,
das weiterhin mindestens einen der folgenden Schritte umfasst:

Bestimmung des zugeordneten Teilnehmerstandortes (1100) aus dem empfangenen erkannten Sprachkanal;

Bestimmung des zugeordneten Teilnehmerstandortes (1100) aus dem empfangenen erkannten Sprachinhalt;

Bestimmung des zugeordneten Teilnehmerstandortes (1100) aus dem erkannten Sprachinhalt und einer Sprechererkennungs-Bibliothek;

Bestimmung des zugeordneten Teilnehmerstandortes (1100) aus dem erkannten Sprachinhalt und einer Spracherkennungs-Bibliothek; und

Bestimmung des zugeordneten Teilnehmerstandortes (1100) aus einer Erkennung innerhalb des Sprachkanals.

Der erteilte Patentanspruch 14 (gemäß EP 1 290 889 B1) in der englischen Fassung unterscheidet sich von Patentanspruch 2 (gemäß EP 1 290 889 B1) nur dadurch, dass der Ausdruck „The method of Claim 1. further comprising at least one of the steps of:“ durch den Ausdruck „The system of Claim 13, wherein said processor (1520) further executes at least one of the steps of:“ ersetzt ist.

Der erteilte Patentanspruch 14 in der deutschen Übersetzung (gemäß EP 1 290 889 B1) unterscheidet sich von Patentanspruch 2 (gemäß EP 1 290 889 B1) nur dadurch, dass der Ausdruck „Verfahren nach Anspruch 1, das weiterhin mindestens einen der folgenden Schritte umfasst:“ durch den Ausdruck „System nach Anspruch 13, wobei der Prozessor (1520) weiterhin mindestens einen der folgenden Schritte ausführt:“ ersetzt ist.

Im Folgenden wird zu jedem der Hilfsanträge 1 bis 8 nur der jeweilige Patentanspruch 1 aufgeführt, da die jeweiligen Patentansprüche 13 dieser Hilfsanträge analoge Änderungen enthalten, und da die jeweiligen Patentansprüche 2 und 14 der Hilfsanträge 1 bis 6 identisch mit den erteilten Patentansprüchen 2 und 14 sind und Merkmale dieser Patentansprüche in die Patentansprüche 1 und 13 nach Hilfsanträgen 7 und 8 aufgenommen worden sind.

Patentanspruch 1 gemäß Hilfsantrag 1 vom 17. Februar 2022 hat folgenden Wortlaut (wobei die Änderungen zum erteilten Patentanspruch 1 gemäß EP 1 290 889 B1 unterstrichen sind und die neuen Merkmale entsprechend bezeichnet sind):

- 1.1 A method of using a back channel
 - 1.1.1 containing a multiplicity of identified speech channels from a multiplicity of user sites (1100)
 - 1.1.2 presented to a speech recognition system (3200) at a wireline node (1300) of a network supporting at least one of cable television delivery and video delivery,
- 1.2 comprising the steps of:
 - 1.2.1 receiving said back channel to create a received back channel,
 - 1.2.2 partitioning said received back channel into a multiplicity of received identified speech channels;
 - 1.2.3 processing each of said multiplicity of said received identified speech channels with said speech recognition system to create corresponding recognized speech content for each received identified speech channel;
 - 1.2.4 responding to said recognized speech content to create a recognized speech content response that is unique for each of said received identified speech channels; and
 - 1.2.5 individually controlling the delivery of entertainment and information services to each user site (1100) in accordance with said recognized speech;
 - 1.2.6 wherein the entertainment and information services comprise a Video-On-Demand service and an Interactive Program Guide.

Patentanspruch 1 gemäß Hilfsantrag 2 vom 17. Februar 2022 hat folgenden Wortlaut (wobei die Änderungen zum erteilten Patentanspruch 1 gemäß EP 1 290 889 B1 unterstrichen sind und die neuen Merkmale entsprechend bezeichnet sind):

- 1.1 A method of using a back channel

- 1.1.1 containing a multiplicity of identified speech channels from a multiplicity of user sites (1100)
- 1.1.2 presented to a speech recognition system (3200) at a wireline node (1300) of a network supporting at least one of cable television delivery and video delivery,
- 1.2 comprising the steps of:
 - 1.2.1 receiving said back channel to create a received back channel,
 - 1.2.2 partitioning said received back channel into a multiplicity of received identified speech channels;
 - 1.2.3 processing each of said multiplicity of said received identified speech channels with said speech recognition system to create corresponding recognized speech content for each received identified speech channel;
 - 1.2.4 responding to said recognized speech content to create a recognized speech content response that is unique for each of said received identified speech channels; and
 - 1.2.5 individually controlling the delivery of entertainment and information services to each user site (1100) in accordance with said recognized speech;
 - 1.2.7 wherein the speech recognition is performed only by said speech recognition system.

Patentanspruch 1 gemäß Hilfsantrag 3 vom 17. Februar 2022 hat folgenden Wortlaut (wobei die Änderungen zum erteilten Patentanspruch 1 gemäß EP 1 290 889 B1 unterstrichen sind und die neuen Merkmale entsprechend bezeichnet sind):

- 1.1 A method of using a back channel
 - 1.1.1 containing a multiplicity of identified speech channels from a multiplicity of user sites (1100)
 - 1.1.2 presented to a speech recognition system (3200) at a wireline node (1300) of a network supporting at least one of cable television delivery and video delivery,
 - 1.2 comprising the steps of:

- 1.2.1 receiving said back channel to create a received back channel,
- 1.2.2 partitioning said received back channel into a multiplicity of received identified speech channels;
- 1.2.3 processing each of said multiplicity of said received identified speech channels with said speech recognition system to create corresponding recognized speech content for each received identified speech channel;
- 1.2.4 responding to said recognized speech content to create a recognized speech content response that is unique for each of said received identified speech channels; and
- 1.2.5 individually controlling the delivery of entertainment and information services to each user site (1100) in accordance with said recognized speech;
- 1.2.8 wherein each user site comprises at least one set top box which is associated with a remote control comprising a microphone and a talk button;
- 1.2.9 and when the talk button is pushed by the user, the remote control sends a talk-button-active command to the set top box;
- 1.2.10 and the set-top box then places an icon on a screen coupled to the set top box and/or otherwise indicates to the user that she or he is listened to.

Patentanspruch 1 gemäß Hilfsantrag 4 vom 17. Februar 2022 hat folgenden Wortlaut (wobei die Änderungen zum erteilten Patentanspruch 1 gemäß EP 1 290 889 B1 unterstrichen sind und die neuen Merkmale entsprechend bezeichnet sind):

- 1.1 A method of using a back channel
 - 1.1.1 containing a multiplicity of identified speech channels from a multiplicity of user sites (1100)
 - 1.1.2 presented to a speech recognition system (3200) at a wireline node (1300) of a network supporting at least one of cable television delivery and video delivery,
- 1.2 comprising the steps of:

- 1.2.1 receiving said back channel to create a received back channel,
- 1.2.2 partitioning said received back channel into a multiplicity of received identified speech channels;
- 1.2.3 processing each of said multiplicity of said received identified speech channels with said speech recognition system to create corresponding recognized speech content for each received identified speech channel;
- 1.2.4 responding to said recognized speech content to create a recognized speech content response that is unique for each of said received identified speech channels; and
- 1.2.5 individually controlling the delivery of entertainment and information services to each user site (1100) in accordance with said recognized speech;
- 1.2.8' wherein each user site contains at least one set top box which is associated with a remote control containing a microphone and a talk button;
- 1.2.11 and wherein upon depressing the talk button on the remote control data are sent to the wireline node alerting the system as to the user site and a potential input.

Patentanspruch 1 gemäß Hilfsantrag 5 vom 17. Februar 2022 hat folgenden Wortlaut (wobei die Änderungen zum erteilten Patentanspruch 1 gemäß EP 1 290 889 B1 unterstrichen sind und die neuen Merkmale entsprechend bezeichnet sind):

- 1.1 A method of using a back channel
 - 1.1.1 containing a multiplicity of identified speech channels from a multiplicity of user sites (1100)
 - 1.1.2 presented to a speech recognition system (3200) at a wireline node (1300) of a network supporting at least one of cable television delivery and video delivery,
- 1.2 comprising the steps of:
 - 1.2.1 receiving said back channel to create a received back channel,

- 1.2.2 partitioning said received back channel into a multiplicity of received identified speech channels;
- 1.2.3 processing each of said multiplicity of said received identified speech channels with said speech recognition system to create corresponding recognized speech content for each received identified speech channel;
- 1.2.4 responding to said recognized speech content to create a recognized speech content response that is unique for each of said received identified speech channels; and
- 1.2.5 individually controlling the delivery of entertainment and information services to each user site (1100) in accordance with said recognized speech;
- 1.2.12 wherein said network comprises a content engine from which said speech recognition system receives content status information;
- 1.2.8' wherein each user site contains at least one set top box which is associated with a remote control containing a microphone and a talk button;
- 1.2.13 wherein the analog signals picked up by the microphone are pre-processed by the remote control;
- 1.2.14 wherein the set top box receives a radio frequency (RF) signal or an infra-red signal from the remote control; and
- 1.2.11 and wherein upon depressing the talk button on the remote control data are sent to the wireline node alerting the system as to the user site and a potential input.

Patentanspruch 1 gemäß Hilfsantrag 6 vom 17. Februar 2022 hat folgenden Wortlaut (wobei die Änderungen zum erteilten Patentanspruch 1 gemäß EP 1 290 889 B1 unterstrichen sind und die neuen Merkmale entsprechend bezeichnet sind):

- 1.1 A method of using a back channel
- 1.1.1 containing a multiplicity of identified speech channels from a multiplicity of user sites (1100)

- 1.1.2 presented to a speech recognition system (3200) at a wireline node (1300) of a network supporting at least one of cable television delivery and video delivery,
- 1.2 comprising the steps of:
 - 1.2.1 receiving said back channel to create a received back channel,
 - 1.2.2 partitioning said received back channel into a multiplicity of received identified speech channels;
 - 1.2.3 processing each of said multiplicity of said received identified speech channels with said speech recognition system to create corresponding recognized speech content for each received identified speech channel;
 - 1.2.4 responding to said recognized speech content to create a recognized speech content response that is unique for each of said received identified speech channels; and
 - 1.2.5 individually controlling the delivery of entertainment and information services to each user site (1100) in accordance with said recognized speech;
 - 1.2.15 wherein the speech recognition system (3200, 1330) communicates with a content engine (1340), indicating a current user location in a menu structure and/or a requested user action, in particular a video request;
 - 1.2.16 and wherein the content engine (1340) provides at least one multimedia stream (1342) to a transceiver (1320) from which it is send to an individual user site.

Patentanspruch 1 gemäß Hilfsantrag 7 vom 17. Februar 2022 hat folgenden Wortlaut (wobei die Änderungen zum erteilten Patentanspruch 1 gemäß EP 1 290 889 B1 unterstrichen sind und die neuen Merkmale entsprechend bezeichnet sind):

- 1.1 A method of using a back channel
 - 1.1.1 containing a multiplicity of identified speech channels from a multiplicity of user sites (1100)

- 1.1.2 presented to a speech recognition system (3200) at a wireline node (1300) of a network supporting at least one of cable television delivery and video delivery,
- 1.2 comprising the steps of:
 - 1.2.1 receiving said back channel to create a received back channel,
 - 1.2.2 partitioning said received back channel into a multiplicity of received identified speech channels;
 - 1.2.3 processing each of said multiplicity of said received identified speech channels with said speech recognition system to create corresponding recognized speech content for each received identified speech channel;
 - 1.2.4 responding to said recognized speech content to create a recognized speech content response that is unique for each of said received identified speech channels; and
 - 1.2.5 individually controlling the delivery of entertainment and information services to each user site (1100) in accordance with said recognized speech;
- 1.2.17 further comprising at least one of the steps of:
 - determining said associated user site (1100) from said recognized speech content;
 - determining said associated user site (1100) from said recognized speech content and a speaker identification library;
 - determining said associated user site (1100) from said recognized speech content and a speech recognition library; and
 - determining said associated user site (1100) from an identification within said speech channel.

Patentanspruch 1 gemäß Hilfsantrag 8 vom 5. Mai 2022 hat folgenden Wortlaut (wobei die Änderungen zum erteilten Patentanspruch 1 gemäß EP 1 290 889 B1 unterstrichen sind und die neuen Merkmale entsprechend bezeichnet sind):

- 1.1 A method of using a back channel

- 1.1.1 containing a multiplicity of identified speech channels from a multiplicity of user sites (1100)
- 1.1.2 presented to a speech recognition system (3200) at a wireline node (1300) of a network supporting at least one of cable television delivery and video delivery,
- 1.2 comprising the steps of:
 - 1.2.1 receiving said back channel to create a received back channel,
 - 1.2.2 partitioning said received back channel into a multiplicity of received identified speech channels;
 - 1.2.3 processing each of said multiplicity of said received identified speech channels with said speech recognition system to create corresponding recognized speech content for each received identified speech channel;
 - 1.2.4 responding to said recognized speech content to create a recognized speech content response that is unique for each of said received identified speech channels; and
 - 1.2.5 individually controlling the delivery of entertainment and information services to each user site (1100) in accordance with said recognized speech;
 - 1.2.17 further comprising at least one of the steps of:
 - determining said associated user site (1100) from said recognized speech content;
 - determining said associated user site (1100) from said recognized speech content and a speaker identification library;
 - determining said associated user site (1100) from said recognized speech content and a speech recognition library.

Wegen der weiteren Einzelheiten wird auf den Akteninhalt verwiesen.

Entscheidungsgründe

Die Klage, mit der die Nichtigkeitsgründe der fehlenden Patentfähigkeit (Art. II § 6 Abs. 1 Satz 1 Nr. 1 IntPatÜG, Art. 138 Abs. 1 lit. a) EPÜ i. V. m. Art. 52, 54 und 56 EPÜ), der unzureichenden Offenbarung (Art. II § 6 Abs. 1 Satz 1 Nr. 2 IntPatÜG, Art. 138 Abs. 1 lit. b) EPÜ i. V. m. Art. 83 EPÜ) und der unzulässigen Erweiterung (Art. II § 6 Abs. 1 Satz 1 Nr. 3 IntPatÜG, Art. 138 Abs. 1 lit. c) EPÜ i. V. m. Art. 123 EPÜ) geltend gemacht werden, ist zulässig. Insbesondere besteht trotz des Erlöschens des Streitpatents durch Zeitablauf am 7. Mai 2021 ein besonderes, eigenes Rechtsschutzbedürfnis der Klägerin, weil die Beklagte unter anderem die Klägerin mit Klageschriftsatz vom 5. Mai 2020 wegen angeblicher Verletzung des Streitpatents vor dem Landgericht Mannheim verklagt hat und deshalb die nicht nur theoretische Gefahr besteht, dass die Klägerin für die Zeit vor Erlöschen des Schutzrechts wegen Verletzung des Patents in Anspruch genommen wird (vgl. *Keukenschrijver*, Patentnichtigkeitsverfahren, 7. Aufl. 2021, Rn. 133 ff. mit umfangreichen Nachweisen zur Rechtsprechung des BGH und BPatG).

Die Klage ist auch begründet. Die angegriffenen Ansprüche 1, 2, 13 und 14 des Streitpatents haben weder in der erteilten Fassung noch in der Fassung eines der Hilfsanträge Bestand.

I.

Die im Schriftsatz der Klägerin vom 27. April 2022 enthaltenen Angriffe auf Basis der Druckschriften **NK18**, **NK19** (WO 97 / 13 368 A1), **NK20** und **NK22** waren trotz Rüge der Beklagten nach § 83 Abs. 4 Satz 1 PatG nicht als verspätet zurückzuweisen.

Damit ist über den Angriff auf das Streitpatent nach den Druckschriften **NK18**, **NK19** (WO 97 / 13 368 A1), **NK20** und **NK22** in der Sache zu entscheiden.

Gemäß § 83 Abs. 4 Satz 1 PatG kann das Patentgericht zwar Angriffsmittel zurückweisen und bei seiner Entscheidung unberücksichtigt lassen. Hierfür ist es aber stets erforderlich, dass Angriffsmittel enthaltender Vortrag tatsächliche oder rechtliche Fragen aufkommen lässt, die in der mündlichen Verhandlung nicht oder nur mit unverhältnismäßigem Aufwand zu klären sind (vgl. Begründung zum Entwurf eines Gesetzes zur Vereinfachung und Modernisierung des Patentrechts, BIPMZ 2009, 307, 315). Kann das an sich verspätete Vorbringen dagegen noch ohne Weiteres in die mündliche Verhandlung einbezogen werden, ohne dass es zu einer Verfahrensverzögerung kommt, liegen die Voraussetzungen für eine Zurückweisung nach § 83 Abs. 4 PatG nicht vor (vgl. *Keukenschrijver*, a. a. O., Rn. 223 mit umfangreichen Nachweisen zur Rechtsprechung des BPatG).

II.

1. Das Streitpatent betrifft die Durchführung einer Spracherkennung an einem Leitungsknoten eines Netzwerks, welches Kabelfernsehen und/oder Video-Verteildienste unterstützt (Streitpatentschrift, Absatz [0001]).

In der Beschreibungseinleitung des Streitpatents ist ausgeführt, dass nur einige wenige Anwendungen (am Anmeldetag) Sprachsteuerungsfunktionen verwendet hätten, die auf den (damals) neuesten Spracherkennungstechnologien basieren. Aufgrund der mittelmäßigen Spracherkennungseffizienz der entsprechenden Spracherkennungssysteme, die typischerweise nur einen begrenzten Befehlsumfang aufwiesen, sei oft ein Sprachtraining nötig gewesen. Daneben sei in High-End-Systemen die maschinelle Verarbeitung natürlicher Sprache eingesetzt worden, welche modernste Software sowie mehrere hundert Megabytes Speicher erfordert hat. Jedoch habe der Stand der Technik die Probleme der Spracherkennung nicht adressiert, die sich an einem zentralisierten Leitungsknoten in einem Netzwerk stellen, das die Lieferung von Videos oder von Kabelfernsehen unterstützt (Streitpatentschrift, Absätze [0002], [0004], [0005]; aus dem Stand der Technik bekannte typische Beispiele für solche Netzwerke sind in den Absätzen [0006] bis [0029] i. V. m. den Figuren 1 und 2 der Streitpatentschrift beschrieben).

Des Weiteren sei es beim Design von Kabelfernseh-Set-Top-Boxen ein wesentliches Ziel gewesen, Informationen von der Kabelfernsehanlage zu den Teilnehmern (d.h. „downstream“) effizient zu übertragen. Hingegen sei die Berücksichtigung der Informationsübermittlung von einem Teilnehmer zur Kabelfernsehanlage („upstream“) viel mehr Beschränkungen unterworfen gewesen.

Da neue Klassen interaktiver Dienste verfügbar geworden seien, habe die effiziente Ausnutzung der Übertragungsbandbreite in Upstream-Richtung an Bedeutung gewonnen (Streitpatentschrift, Absatz [0013]).

Ferner seien umfangreiche Forschungen zu den Mechanismen der Spracherkennung durchgeführt worden; die dabei gemachten Fortschritte hätten es Börsenmaklern erlaubt, mittels ihrer Tischcomputer sprachbasiert zu handeln (Streitpatentschrift, Absatz [0030]).

Jedoch seien mehrere zentrale Fragestellungen nicht geklärt worden, die für Kabelfernsehen, Videoverteilssysteme und den Handel von großer Wichtigkeit seien. So habe es kein System gegeben, bei dem Teilnehmer basierend auf einer Spracherkennung identifiziert werden, die über ein Netzwerk hinweg ausgeführt wird, das Kabelfernsehen und/oder Video-Verteildienste unterstützt. Auch sei kein System vorhanden gewesen, das über ein Kabelfernseh- und/oder Video-Verteilnetzwerk durchgeführte Echtzeit-Auktionen und Vertragsabschlüsse in ausreichendem Maße unterstützt und auf der Identifikation der Teilnehmer durch Spracherkennung beruht (Streitpatentschrift, Absatz [0031]).

2. Eine Aufgabe wird im Streitpatent nicht ausdrücklich genannt. Jedoch ist aus den oben genannten Absätzen der Streitpatentschrift sowie aus Patentanspruch 1 und 13 die **Aufgabe** abzuleiten, ein Verfahren bzw. ein System anzugeben, das eine Spracherkennung für mehrere Teilnehmer über ein Kabelfernseh- und/oder Videoverteilnetzwerk hinweg ermöglicht und dadurch eine individuelle Kontrolle der Lieferung von Unterhaltungs- und Informationsdiensten gewährleistet.

3. Als **Fachmann**, der mit der Lösung dieser Aufgabe betraut wird, ist ein Elektrotechnikingenieur der Fachrichtung Nachrichtentechnik anzusehen, der mehrjährige Erfahrung in der Konzeption und Entwicklung von Kabelfernseh- und Videoverteilnetzwerken inklusive der zugehörigen Benutzerschnittstellen besitzt und mit den jeweils gültigen Standards zum Betrieb solcher Netzwerke vertraut ist.

4. Dieser Fachmann legt den Merkmalen der von der Nichtigkeitsklage angegriffenen Patentansprüche 1, 2, 13 und 14 folgendes Verständnis zugrunde:

4.1 Das Streitpatent definiert nicht ausdrücklich, was unter einem Rückkanal sowie unter einem Sprachkanal zu verstehen sein soll, der eine Vielzahl von erkannten Sprachkanälen von einer Vielzahl von Teilnehmerstandorten enthält (vgl. Merkmale 1.1 und 1.1.1 sowie 13.2 bis 13.4).

Aus Sicht des Fachmanns ist ein Rückkanal ein Kommunikationskanal oder Übertragungsweg, auf dem Signale (z.B. Sprachsignale) von den Netzwerkteilnehmern an eine Sendestation (z.B. an eine Kopfstation („Headend“) eines Kabelfernsehnetzwerks) „upstream“ übermittelt werden, d.h. entgegen der Übertragungsrichtung, die beim Transfer der von der Sendestation bereitgestellten Inhalte an die Teilnehmer verwendet wird.

Dementsprechend ist ein Sprachkanal ein Kommunikationskanal oder Übertragungsweg eines Kommunikationsnetzwerks, auf dem Sprachinformationen übertragen werden.

Darüber hinaus sind aber auch die auf einem Rückkanal übermittelten Informationssignale oder Informationen selbst als Rückkanal anzusehen (bzw. als Sprachkanal, falls diese auch Sprachinformationen umfassen). So ist dem Streitpatent beispielsweise zu entnehmen, dass ein Rückkanal in einer Uplink-Kommunikation enthalten sein, empfangen und an einen Spracherkennungsrechner

geliefert werden kann, und dass ein Spracherkennungssystem erkannte Sprachkanäle verarbeitet (Streitpatentschrift, Absätze [0163], [0164], [0166], [0258], [0263]; s. auch Merkmale **1.2.1**, **1.2.3**, **13.3**, **13.5.1**).

4.2 Die in Merkmal **1.1.1** genannten „identified speech channels“ sind Sprachkanäle, die entsprechend dem Sinngehalt des Begriffs „identified“ insbesondere erkannt, festgelegt, ermittelt, bestimmt oder gekennzeichnet worden sind; sie werden nach dem Empfang des Rückkanals (vgl. Merkmale **1.2.1**, **13.3**) in „received identified speech channels“ aufgeteilt (Merkmale **1.2.2** und **13.4**) und liegen bereits vor, bevor die eigentliche Spracherkennung gemäß den Merkmalen **1.2.3** und **13.5.1** durchgeführt worden ist.

4.3 Ein Leitungsknoten („wireline node“, vgl. Merkmale **1.1.2**, **13.2**) ist vor dem Hintergrund von Absatz [0005] der Streitpatentschrift insbesondere als ein Netzwerkknoten anzusehen, der Video- oder Kabelfernsehdienste für mehrere Teilnehmer über physische Leitungen (z.B. Kabel) erhält und/oder bereitstellt, wie dies beispielsweise bei einem Headend eines Kabelfernsehtzwerks der Fall sein kann (vgl. Streitpatentschrift, Figur 1 i. V. m. Absätzen [0008], [0012], [0016]).

4.4 Ein Aufteilen des empfangenen Rückkanals in eine Vielzahl empfangener erkannter Sprachkanäle („partitioning said received back channel into a multiplicity of received identified speech channels“, vgl. Merkmale **1.2.2** und **13.4**) interpretiert der Fachmann derart, dass Sprachinformationen bzw. Sprachinformationssignale, die auf dem Rückkanal als Datenstrom übertragen und empfangen worden sind, in einzelne Sprachinformationsabschnitte aufgetrennt, segmentiert oder zerlegt oder in verschiedene Gruppen eingeteilt werden.

4.5 Was unter einem erkannten Sprachinhalt („recognized speech content“, vgl. Merkmale **1.2.3** und **13.5.1**) zu verstehen sein soll, der aus der Verarbeitung der erkannten Sprachkanäle resultiert, erläutert das Streitpatent nicht. Der Beschreibung des Streitpatents ist in diesem Zusammenhang zu entnehmen, dass zur Sprachverarbeitung eine Grammatik verwendet wird (Absatz [0098]) und als Spracherkennungsergebnisse Worte, Anfragen, Textstrings oder Phrasen

ausgegeben werden (vgl. Absätze [0036] bis [0038], [0041], [0042] - „Thus, when a command is spoken and recognized the system returns the key word“; Absätze [0088], [0112] - „recognition of a spoken request“; Absätze [0111], [0116] bis [0119] - „recognized text string“, „By displaying the text of the possible recognition results [...]“, „recognized phrase“).

Der Fachmann versteht daher unter dem „erkannten Sprachinhalt“ Informationen, die bereits in den an das Spracherkennungssystem übertragenen Sprachsignalen enthalten sind, und die von diesem ermittelt und in einer bestimmten Repräsentation - z.B. als Wort, Wortfolge oder Textstring - ausgegeben werden.

Aus Sicht des Fachmanns ist dieser „erkannte Sprachinhalt“ mit dem in den Merkmalen **1.2.5** und **13.5.3** genannten Begriff „erkannte Sprache“ („recognized speech“) zu identifizieren, da sich beide Begriffe auf das Ergebnis einer Sprachverarbeitung durch ein Spracherkennungssystem beziehen und eine solche Verarbeitung nur in Merkmal **1.2.3** bzw. **13.5.1** des erteilten Patentanspruchs 1 bzw. 13 vorgenommen wird.

Ferner ist festzuhalten, dass sämtliche Spracherkennungsoperationen, auf die Merkmal **1.2.3** Bezug nimmt, gemäß dem Wortlaut des erteilten Patentanspruchs 1 von dem in Merkmal **1.1.2** genannten Spracherkennungssystem an dem Leitungsknoten ausgeführt werden (vgl. Merkmal **1.2.3**: „processing [...] with said speech recognition system to create corresponding recognized speech content“) - denn von einem anderen Spracherkennungssystem ist im erteilten Patentanspruch 1 nicht die Rede.

Im Übrigen lässt das Streitpatent auch nicht darauf schließen, dass ein bestimmter erkannter Sprachinhalt zusätzlich auch noch von einem (weiteren) Spracherkennungssystem ermittelt wird, das sich außerhalb des Leitungsknotens befindet, der die Sprachverarbeitung und -erkennung gemäß den Merkmalen **1.1.2** und **1.2.3** ausführt. Zwar finden gemäß den Absätzen [0064], [0065] und [0091] Vorverarbeitungsoperationen an dem Teilnehmerstandort statt, aber das Streitpatent subsumiert diese Operationen nicht unter eine Spracherkennung (vgl.

Absätze [0097] bis [0111] - der Sprachprozessor der Speech-Engine, der den im ersten Satz von Absatz [0108] angesprochenen ersten Spracherkennungsschritt sowie die sich gemäß Absatz [0109] daran anschließende, in Absatz [0111] beschriebene Spracherkennung ausführt, befindet sich an dem Leitungsknoten). Auch bei dem in Figur 21 gezeigten Netzwerk wird die Sprachverarbeitung für die Headends 100 und 106 und die Knoten 120 und 124 jeweils von genau einem der mit den Bezugszeichen 1410 gekennzeichneten „augmentierten“ Headends ausgeführt; für das in Figur 22 gezeigte Netzwerk gilt Entsprechendes (Streitpatentschrift, Absätze [0235], [0242]).

Somit ist Merkmal **1.2.3** derart zu verstehen, dass die darin genannten Spracherkennungsoperationen ausschließlich von dem mit Merkmal **1.1.2** beanspruchten Spracherkennungssystem an dem Leitungsknoten ausgeführt werden.

4.6 Ein Unterhaltungs- bzw. ein Informationsdienst (vgl. Merkmale **1.2.5**, **13.5.3**) ist aus fachmännischer Sicht zunächst einmal eine Funktionalität eines Kommunikationsnetzwerks, die der Bereitstellung und Übermittlung bestimmter Daten (z.B. von Videos, Filmen, Fernsehprogrammen, Informationen zu einem Online-Einkauf oder Internetinhalten) zur Unterhaltung bzw. zur Information von Netzwerkteilnehmern dient (vgl. Streitpatentschrift, Absatz [0035]).

Eine Lieferung von Unterhaltungs- und Informationsdiensten („delivery of entertainment and information services“, vgl. Merkmale **1.2.5**, **13.5.3**) umfasst aus fachmännischer Sicht sowohl die Übermittlung als auch dem Empfang von Daten an einen bzw. einem Teilnehmerstandort durch mindestens einen Unterhaltungs- und mindestens einen Informationsdienst. Diese Daten können z.B. im Fall eines Textdienstes Nachrichtenartikel sein (vgl. **NK9**, Spalte 5, Zeile 46 und 47 - „text services such as news articles [...]“), aber auch Videos oder Programmführerinformationen, falls der Dienst ein Video-on-Demand-Dienst ist oder ein elektronischer Programmführer mittels des Dienstes übertragen wird. Bei einer Lieferung eines Unterhaltungs- bzw. Informationsdienstes können nach fachmännischem Verständnis auch Daten übermittelt oder empfangen werden, die

ein Programm repräsentieren, das benötigt wird, um einen solchen Dienst ausführen zu können (z.B. eine elektronische Programmführersoftware, die auf einer Set-Top-Box installiert werden soll).

Eine individuelle Kontrolle der Lieferung eines Dienstes an jeden Teilnehmerstandort („individually controlling the delivery [...] to each user site“, vgl. Merkmale **1.2.5**, **13.5.3**) in Übereinstimmung mit der erkannten Sprache liegt etwa dann vor, wenn jeder einzelne Teilnehmer durch seine mündlichen Äußerungen die Art der Daten oder deren Bedeutungsinhalt bestimmt, die durch den Dienst an seinen Standort übermittelt werden, oder wenn jeder Teilnehmer beeinflusst, wie diese Daten an seinen Standort und/oder die Standorte der anderen Teilnehmer gesendet oder dort empfangen werden.

4.7 Ein Prozessor gemäß den Merkmalen **13.5** bis **13.5.3** ist eine Komponente, die Daten verarbeitet und Mittel umfasst, die den in den Merkmalen **13.5.1** bis **13.5.3** aufgezählten Zwecken dienen (vgl. auch BGH, Urteil vom 3. August 2021, X ZR 71/19, juris und GRUR 2021, 1375 – *Bediengerät für Spiele*, Rn. 79).

4.8 Unter einem „genannten zugeordneten Teilnehmerstandort“ („said associated user site“) im Sinne der Patentansprüche 2 und 14 des Streitpatents versteht der Fachmann einen der in Merkmal **1.1.1** bzw. **13.2** genannten Teilnehmerstandorte, da die Patentansprüche 1 bzw. 13 keine anderen Merkmale enthalten, die sich auf einen Teilnehmerstandort beziehen. Ein Teilnehmerstandort kann ein Raumbereich sein, der eine Set-Top-Box enthält oder mittels einer Raumnummer oder Adressangabe identifiziert werden kann (vgl. Streitpatentschrift, Absatz [0007] - „Each user site contains a Set Top Box“; Absatz [0184] - „[...] may identify the user site as Room 432 or 10 Main Street“), oder auch eine Set-Top-Box selbst (vgl. Absatz [0258] - „speech channels from multiple user sites (STBs) 1100“; Absatz [0063] - „A given residence may include more than one set-top box 1100, each of which has a distinct address in the network [...] Each constitutes a distinct user site“).

III.

Das Streitpatent hat in der erteilten Fassung keinen Bestand, weil die jeweiligen Gegenstände der angegriffenen Patentansprüche 1, 2, 13 und 14 nicht patentfähig sind.

1. Die Lehre des erteilten **Patentanspruchs 1** beruht gegenüber dem der Druckschrift **NK11** entnehmbaren Stand der Technik auf keiner erfinderischen Tätigkeit.

1.1 Die Druckschrift **NK11** betrifft interaktive Fernsehsysteme und stellt Mittel und Verfahren vor, wie ein Server verwendet werden kann, um Informationen, die von einem Teilnehmer zur Verfügung gestellt werden, in ein Datenformat zu konvertieren, das von einer interaktiven Fernsichtwendung verwendet werden kann (**NK11**, Seite 1, Zeile 8 bis 10).

In der Einleitung der **NK11** ist ausgeführt, dass interaktive Fernsehsysteme Nutzerinteraktionen unterstützen und es daher den Fernsehteilnehmern ermöglichen, eine Vielzahl von Produkten und Diensten zu bestellen, Informationen zu bestimmten Fernsehprogrammen anzufordern oder elektronische Nachrichten zu versenden (Seite 1, Zeile 13 bis 18).

Konkret ist ein System zur Verteilung interaktiver Fernsichtwendungen und Fernsehprogramme an eine Reihe von Teilnehmern bzw. Zuschauern beschrieben, wobei die Sendungen insbesondere auch über ein Kabelfernsehnetzwerk - also kabelgebunden – an deren Empfangsstationen übertragen werden können (Figur 1 i. V. m. Seite 5, Zeilen 1 und 2 - „a system for distribution of interactive television programs [...] to a series of viewers“ sowie Seite 5, Zeile 27 bis 28 - „any broadcast medium (e.g., CATV [...]) may be used“; „CATV“ = „cable television“). Das System erlaubt es einem Teilnehmer, digitalisierte Bilddaten eines Grafiktablets (Seite 4, Zeile 30 bis 32; Claims 18, 20 und 21) oder Sprachdaten, die auf den von einem Mikrophon erfassten mündlichen Äußerungen des Teilnehmers basieren (Seite 2, Zeile 15 bis 18, Seite 3, Zeile 20 bis 24; Seite 10, Zeile 20 bis 27;

Claims 18 und 22), über seine Set-Top-Box 22 als „nicht-textuelle“ Daten an einen an einer Sendestation 10 angeordneten Server 13 zu übertragen. Dieser verwendet eine Spracherkennungssoftware, um die Sprachdaten in Textdaten zu konvertieren, die anschließend an die Set-Top-Box eines jeweiligen Teilnehmers zurückgeschickt und dort angezeigt werden (Seite 3, Zeilen 20 bis 24; Seite 5, Zeilen 6 bis 8; Seite 10, Zeilen 20 bis 27). Als Rückkanal („return path“) zur Übertragung der Sprachdaten an die Sendestation wird insbesondere ein Datenübertragungsweg verwendet, der einen Teil der Bandbreite eines Rundfunkkanals umfasst und von den Set-Top-Boxen der Teilnehmer über das Kabelfernsehnetzwerk und den Demultiplexer 18 zu dem Server 13 hin verläuft (Claims 1, 4 und 6 - „wherein said return path comprises a portion of the bandwidth of the broadcast channel“; Figur 1 i. V. m. Seite 5, Zeile 27 und 28 sowie Seite 8, Zeilen 13 bis 16).

Somit zeigt die Druckschrift **NK11** ein Verfahren zur Verwendung eines Rückkanals (Merkmal **1.1**).

Es ist selbstverständlich, dass in einem Kabelfernsehnetzwerk grundlegende Benutzerschnittstellenfunktionen nicht nur für einen einzigen, sondern für eine große Anzahl von Nutzern implementiert werden. Dies bedeutet im vorliegenden Fall, dass die Sprachdaten mehrerer Teilnehmer über den Rundfunk-Rückkanal gesendet werden, so dass dieser auch eine Vielzahl von Sprachkanälen von einer Vielzahl von Teilnehmerstandorten enthält. Dass die Sprachdaten vor ihrer Übermittlung festgelegt worden sein müssen und somit „identified speech channels“ im Sinne von Merkmal **1.1.1** sind, ist selbstverständlich.

Der Server 13, der insbesondere die Spracherkennung ausführt, befindet sich an der Sendestation 10, die ein kabelgebundener Leitungsknoten eines Kabelfernsehnetzwerks ist (vgl. Figur 1 sowie Seite 5, Zeilen 6 und 7 sowie Zeilen 27 und 28; Merkmal **1.1.2**).

Die Sprachdaten, die auf dem Rückkanal an den Server 13 übertragen worden sind, werden durch eine Spracherkennungssoftware in Textform konvertiert (Seite 10, Zeilen 25 und 26); dazu müssen sie selbstverständlich zuvor netzwerkseitig

empfangen worden sein (Merkmale **1.2**, **1.2.1** und **1.2.3**). Als Reaktion auf die Spracherkennung kann der von dem Spracherkennungssystem ermittelte Text - dieser repräsentiert den Inhalt der gesprochenen Äußerungen des Teilnehmers - an die Set-Top-Boxen der Teilnehmer geschickt und dort angezeigt oder von beliebigen interaktiven Anwendungen verwendet werden (vgl. Abstract; Claims 1 und 2; Seite 2, Zeilen 33 und 34; Seite 10, Zeile 26; Merkmal **1.2.4**).

Weiterhin zeigt die Druckschrift **NK11**, dass ein Demultiplexer 18 in dem Rundfunk-Rückkanal des in Figur 1 dargestellten Systems angeordnet ist. Dem Fachmann ist bewusst, dass auf einem solchen Kanal Sprachinformationen übertragen werden, indem Zeitmultiplexverfahren wie beispielsweise TDMA-Verfahren („TDMA“ = „Time Division Multiple Access“) zum Einsatz kommen. Solche Verfahren hatten bereits vor den beiden Prioritätszeitpunkten Einzug in einen europäischen Telekommunikationsstandard gefunden; dabei werden Sprachdatenpakete verschiedener Teilnehmer in aufeinanderfolgenden Zeitschlitzen vermischt („multiplext“) auf demselben Kanal gesendet und anschließend wieder aufgeteilt („demultiplext“; vgl. **NK16**, Abschnitt 4.2, Seite 9 - „Interaction Channel (IC): A bi-directional IC is established between the service provider and the user for interaction purposes. It is formed by [...] Return Interaction path [...] Also commonly known as return channel“; Abschnitt 5 i. V. m. Seiten 10 bis 13, insbesondere Figuren 2 und 3, Abschnitt 5.1, erster und zweiter Absatz sowie Abschnitt 5.1.3, insbesondere vierter, vorletzter und letzter Absatz - „Within upstream channels, users send packets with TDMA type access. This means that each channel is shared by many different users“; „There are different access modes for the upstream slots: [...] reserved slots with fixed rate reservation [...] e.g., for voice, audio“ [...] These slots may be mixed on a single carrier [...]).

Der Fachmann wird daher erkennen, dass der Demultiplexer 18 der Aufteilung der Sprachdatenpakete verschiedener Teilnehmer dient, die im Rahmen eines solchen TDMA-Verfahrens auf dem Rundfunk-Rückkanal des Kabelfernsehsystems zu dem Server 13 gesendet worden sind. Von anderen Daten, die auf diesem Rückkanal übertragen werden, ist in **NK11** auch nicht die Rede.

Somit entnimmt der Fachmann unter Zuhilfenahme seines Fachwissens der Druckschrift **NK11** auch das Merkmal **1.2.2**.

Bereits das Zurücksenden der erkannten Texte an die Set-Top-Boxen und das Bereitstellen dieser Texte zur Verwendung in einer bestimmten interaktiven Anwendung kann als eine individuelle Kontrolle der Lieferung eines jeweiligen Informationsdienstes im Sinne von Merkmal **1.2.5** angesehen werden. Denn dabei beeinflussen die einzelnen Teilnehmer die Textdaten, die von dem Spracherkennungssystem ermittelt und an die Teilnehmer transferiert werden, individuell durch ihre sprachlichen Äußerungen. **NK11** führt mehrere solche interaktiven Anwendungen im Kontext einer Handschriftenerkennung an, nämlich einen E-Mail-, einen Fax- und einen elektronischen Einkaufsdienst (Seite 9, Zeile 23 bis Seite 10, Zeile 19). Dass diese Anwendungen anstelle handschriftlicher Texteingaben genauso gut auch gesprochene Äußerungen der Teilnehmer verarbeiten können, ist dem Fachmann aufgrund der Claims 1, 7 und 11 sowie 18 bis 22 der **NK11** klar, denn diese sehen Sprach- und Handschriftenerkennung jeweils als Alternativen vor.

Somit entnimmt der Fachmann der Druckschrift **NK11** zumindest eine „individuelle Kontrolle der Lieferung von Informationsdiensten in Übereinstimmung mit der erkannten Sprache zu jedem Teilnehmerstandort“ (**Teilmerkmal** von Merkmal **1.2.5**).

Eine „individuelle Kontrolle der Lieferung von Unterhaltungsdiensten in Übereinstimmung mit der erkannten Sprache zu jedem Teilnehmerstandort“ (**restliches Teilmerkmal** von Merkmal **1.2.5**) geht allerdings nicht unmittelbar aus **NK11** hervor.

1.2 Der Fachmann gelangt auch zu dem **restlichen Teilmerkmal** von Merkmal **1.2.5**, ohne erfinderisch tätig zu werden.

So kann das in Druckschrift **NK11** beschriebene System in etlichen interaktiven Anwendungen verwendet werden, die auf einer Set-Top-Box ablaufen (Seite 9,

Zeile 23 i. V. m. Claim 1 und Seite 2, Zeile 33 und 34). Das bedeutet, dass diese Anwendungen die von dem Server 13 bereitgestellten Textdaten verwenden.

Aus **NK11** geht ferner hervor, dass ein Teilnehmer mittels solcher interaktiver Anwendungen Informationen zu bestimmten Programmen anfordern oder im Rahmen eines elektronischen Einkaufsdienstes Produkte oder Dienste kaufen kann, indem er entsprechende Menüeinträge auswählt (Seite 1, Zeile 17 und 18; Seite 1, Zeile 38 bis Seite 2, Zeile 5; Spalte 10, Zeile 2 und 3 sowie Zeile 14 bis 19). Diese Produkte oder Dienste können selbstverständlich auch Datenfolgen - z.B. Videos oder elektronische Zeitschriftenartikel - sein, die dem Teilnehmer zu Informations- und Unterhaltungszwecken über das Kabelfernsehnetzwerk elektronisch übermittelt werden.

Für den Fachmann liegt es daher auf der Hand, die von dem Server 13 bereitgestellten Textdaten auch in interaktiven Einkaufsdienstanwendungen zur menügesteuerten Auswahl von elektronisch übermittelten Produkten oder Diensten im Rahmen einer über die Set-Top-Box ausgeführten Kauftransaktion zu verwenden. Das bedeutet, dass ein jeweiliger Teilnehmer durch seine mündlichen Äußerungen die Art und den Bedeutungsinhalt der gekauften Datenfolgen bestimmt, sodass dabei insbesondere eine „individuelle Kontrolle der Lieferung von Unterhaltungsdiensten in Übereinstimmung mit der erkannten Sprache zu jedem Teilnehmerstandort“ gemäß dem **restlichen Teilmerkmal** von Merkmal **1.2.5** vorgenommen wird (s.o., Abschnitt **II.4.6**).

1.3 Die Beklagte argumentiert, die **NK11** offenbare keinen Rückkanal, der mehrere Kanäle umfasst und an einem netzwerkseitigen Spracherkennungssystem eingegeben wird. Der Fachmann könne allenfalls mitlesen, dass der DEMUX-Block 18 funktional zum MUX-Block 17 gehört und zur Trennung eines Sprachsignals von üblichen rückwärtsgerichteten Kontrollsignalen (z.B. ACK/NACK-Signalen) anderer „gemuxter“ Sendequellen (wie etwa der „TV Program Source 11“ und der „Application Source 12“) verwendet werden kann, und dass die im Downstream-Rundfunkkanal gesendeten Rücksignale (z.B. der „Application Source 12“) durch

den DEMUX-Block 18 vom Sprachsignal mit der zu transkribierenden Information getrennt werden müssen.

Des Weiteren werde in **NK11** allein eine Transkription vorgenommen, d.h. eine Transformation einer gesprochenen Sprache in eine textualisierte Sprache bzw. eine „blinde“ Übersetzung der nicht-textuellen Information in eine textuelle Information, aber keine Spracherkennung, die sich auf einen anspruchsgemäßen Sprachinhalt richtet.

Es sei auch nicht offenbart, eine Informationsbereitstellung auf Basis der Sprachdaten zu einer Vielzahl von Nutzerstandorten netzwerkseitig zu steuern. Eine netzwerkseitige Kontrolle mit einer allein für die Set-Top-Box sinngemäßen Information sei nicht möglich, da der Server 13 keine Kopplung zur „TV Program Source 11“ und zur „Application Source 12“ aufweise, so dass die Sprachübertragung als autonomer, geschlossener, vom TV-Netzwerk unabhängiger Kreis offenbart sei. Ein Anlass, eine netzwerkseitige Kontrolle in den Blick zu nehmen, ergebe sich für den Fachmann nicht.

Diese Argumente vermögen nicht zu überzeugen.

Zwar ist der Beklagten insoweit zuzustimmen, dass die **NK11** nicht ausdrücklich davon spricht, dass ein mehrere Sprachkanäle umfassender Rückkanal in ein netzwerkseitiges Spracherkennungssystem eingegeben wird.

Jedoch ist es unrealistisch, dass in einem Kabelfernsehsystem Benutzerschnittstellenfunktionalitäten - wie z.B. die Eingabe von gesprochener Sprache über eine Fernbedienung oder die Anzeige der von einem Spracherkennungssystem ermittelten Texte - ausschließlich für einen einzigen Teilnehmer vorgesehen sein sollen. Dass die in **NK11** dargestellte Erfindung für mehrere Empfangsstationen mehrerer Teilnehmer konzipiert ist, ergibt sich bereits anhand der Hinweise auf Seite 5, Zeile 1, 2 und Zeile 22 bis 24 sowie aus dem Umstand, dass ein Server (hier: der Spracherkennungsserver 13) immer mehrere „Clients“ (hier: mehrere Set-Top-Boxen) bedient. Wenn der in **NK11** beschriebene

Rundfunk-Rückkanal zum Einsatz kommt, müssen also die Sprachdaten mehrerer Teilnehmer über diesen Kanal übertragen werden. Die Erkenntnis, dass diese Übertragung üblicherweise mittels eines Zeitmultiplex-Verfahrens bewerkstelligt werden kann, bei dem einzelne Sprachdatenpakete netzwerkseitig aufgeteilt („demultiplext“) werden, wofür sich gerade der Demultiplexer 18 anbietet, beruht lediglich auf dem Fachwissen des Fachmanns.

Es ist auch nicht plausibel, dass der DEMUX-Block 18 zur Trennung eines (einzigen) Sprachsignals von rückwärtsgerichteten Kontrollsignalen der „TV Program Source 11“ oder der „Application Source 12“ verwendet wird. Denn dafür, dass solche Kontrollsignale von diesen beiden Quellen gesendet werden, liefert die **NK11** keinerlei Anhaltspunkt; zudem sind Kontrollsignale wie z.B. die von der Beklagten angeführten ACK/NACK-Bestätigungssignale („ACK“/„NACK“ = „acknowledgement“/„negative-acknowledgement“) üblicherweise an die sendende Einheit - also hier eine der Quellen 11 oder 12 - gerichtet, wohingegen das Signal, das den DEMUX-Block 18 gemäß Figur 1 der **NK11** verlässt, an den Server 13 gesendet wird. Dieser ist gemäß Figur 1 gerade nicht mit den Quellen 11 oder 12 verbunden.

Des Weiteren sind die Texte, die von dem Spracherkennungssystem der **NK11** ausgegeben werden, durchaus als erkannte Sprachinhalte anzusehen, da sie die gesprochenen Äußerungen der Teilnehmer, die in den über den Rückkanal an das Spracherkennungssystem übertragenen „nicht-textuellen“ Daten enthalten sind, in einer anderen - einer „textuellen“ - Repräsentation darstellen. Zudem ist nicht erkennbar, dass die beanspruchte Spracherkennung über die gemäß **NK11** vorgenommene Spracherkennung hinausgeht, da diese beiden Spracherkennungsverfahren Worte und Textstrings ausgeben (s.o., Abschnitt **II.4.5** bzw. **NK11**, Seite 3, Zeile 23 - „textual data“).

Ferner trifft es zwar zu, dass **NK11** nicht unmittelbar zeigt, dass die Quellen 11 und 12 auf Basis der erkannten Sprachinhalte gesteuert werden. Jedoch ist damit ein Vorliegen von Merkmal **1.2.5** noch nicht ausgeschlossen. Denn ein Teilnehmer bestimmt gemäß **NK11** durch seine individuellen Äußerungen, die er im Rahmen

der Ausführung einer jeweiligen interaktiven Anwendung tätig, sowohl den Bedeutungsinhalt als auch den Übermittlungszeitpunkt der Textdaten, die von dem Server 13 an seinen Standort übermittelt werden. Dadurch beeinflusst er die netzwerkseitige Bereitstellung der Informationen, die an ihn im Rahmen des Dienstes übertragen werden, der der interaktiven Anwendung entspricht (s.o., Abschnitt **II.4.6**). Da die bereitgestellten Textdaten naturgemäß von den gesprochenen Äußerungen abhängen, erfolgt die Informationsbereitstellung auch in Übereinstimmung mit dem erkannten Sprachinhalt, wie mit Merkmal **1.2.5** beansprucht. Zudem liefert die **NK11** - wie oben dargelegt - auch einen konkreten Anlass, die Lieferung von elektronisch übermittelten Unterhaltungs- und Informationsdiensten netzwerkseitig zu beeinflussen.

2. Auch der Gegenstand des erteilten **Patentanspruchs 13** beruht auf keiner erfinderischen Tätigkeit.

Die Merkmale des Patentanspruchs 13 des Streitpatents gehen nur insoweit über übliche Mittel hinaus, die in einem System zur Durchführung des Verfahrens nach dem erteilten Patentanspruch 1 zwangsläufig vorhanden sein müssen, als dass ein und derselbe Prozessor für die in den Merkmalen **13.5.1** bis **13.5.3** beschriebenen Zwecke verwendbar sein muss.

Wenn die in Figur 1 der **NK11** gezeigten Stationen 10 und 20 wie oben in Abschnitt **III.1.1** beschrieben eingesetzt werden, bilden sie eine Komponente, die Daten verarbeitet und Mittel umfasst, die den mit den Merkmalen **13.5.1** bis **13.5.3** beanspruchten Zwecken dienen. Diese Komponente kann als ein Prozessor im Sinne der Merkmale **13.5** bis **13.5.3** angesehen werden (s.o., Abschnitt **II.4.7**).

Somit enthält der Patentanspruch 13 nichts, womit eine erfinderische Tätigkeit begründet werden könnte.

3. Ausgehend von der Lehre der Druckschrift **NK11** war damit der jeweilige Gegenstand der erteilten Patentansprüche 1 und 13 für den Fachmann bereits vor beiden Prioritätszeitpunkten nahegelegt.

4. Die auf den Patentanspruch 1 bzw. den Patentanspruch 13 rückbezogenen, von der Nichtigkeitsklage ebenfalls angegriffenen Patentansprüche 2 und 14 des Streitpatents enthalten nichts, was eine Patentfähigkeit rechtfertigen könnte.

4.1 Patentanspruch 2 fügt der Lehre von Patentanspruch 1 mindestens einen der folgenden fünf Teilschritte hinzu:

determining said associated user site (1100) from said received identified speech channel;

determining said associated user site (1100) from said recognized speech content;

determining said associated user site (1100) from said recognized speech content and a speaker identification library;

determining said associated user site (1100) from said recognized speech content and a speech recognition library; and

determining said associated user site (1100) from an identification within said speech channel.

Die einzelnen Teilschritte beschreiben verschiedene Möglichkeiten zur Bestimmung, Festlegung oder Ermittlung des „zugeordneten“ Teilnehmerstandorts („determining said associated user site from [...]“). Zwar nimmt der erteilte Patentanspruch 1 nicht auf einen „zugeordneten“ Teilnehmerstandort Bezug, jedoch ist aus dem Gesamtzusammenhang der streitpatentgemäßen Lehre unmittelbar ersichtlich, dass unter einem solchen Standort einer der in den Merkmal **1.1.1** und **1.2.5** genannten Teilnehmerstandorte („user sites“) zu verstehen sein muss.

Unter einer "identification within said speech channel" gemäß dem fünften Teilschritt versteht der Fachmann insbesondere eine Kennzeichnung, die zusammen mit den Sprachinformationen mindestens eines Teilnehmers auf einem Sprachkanal übertragen wird. Auch ein Rundfunk-Rückkanal eines Kabelfernsehsystems, auf dem Sprachinformationen übertragen werden, bildet einen solchen Sprachkanal.

4.2 Die Merkmale des Patentanspruchs 2 können keine erfinderische Tätigkeit begründen.

4.2.1 So wird sich der Fachmann ausgehend von der Lehre der Druckschrift **NK11** zur Implementierung der Übertragung der Sprachdatenpakete auf dem Rundfunk-Rückkanal insbesondere auf etablierte Vorgehensweisen zur technischen Realisierung des Rundfunk-Rückkanals stützen, wie sie etwa in gängigen Telekommunikationsstandards zum Einsatz kommen. In diesem Zusammenhang findet er in Druckschrift **NK16**, einer Beschreibung der Implementierung des Interaktionskanals für Kabelfernsehsysteme gemäß dem ETS-Standard, den Hinweis, dass die MAC- und NSAP-Adressen der Set-Top-Boxen - also bestimmte Kennzeichnungen - herangezogen werden, um die von verschiedenen Teilnehmern auf dem Rückkanal übertragene Informationen netzwerkseitig zu unterscheiden (**NK16**, Abschnitt 5.1.3, erster und zweiter Absatz - „Two addresses are stored in STBs in order to identify users on the network: MAC address [...] NSAP address“, „Upstream information may come from any user in the network and shall therefore also be differentiated at the INA using the set of addresses defined above“; „INA“ = „interactive network adapter“, vgl. Seite 7 unten sowie Figur 2).

Für den Fachmann liegt es daher auf der Hand, die MAC- und/oder NSAP-Adresse(n) auf dem Rückkanal zusammen mit den Sprachdatenpaketen des Teilnehmers - z.B. als Bestandteile der Paket-Header - an die Sendestation 10 zu übertragen, so dass die Sprachdaten der einzelnen Teilnehmer durch diejenigen Einheiten der Sendestation 10 unterschieden werden können, die zur Verarbeitung der Sprachdatenpakete vorgesehen sind.

Im Übrigen muss auch in der Sendestation 10 eine Information darüber vorhanden sein, an welchen Teilnehmer die Ergebnisse eines jeweiligen Spracherkennungsvorgangs zu übermitteln sind. Dazu ist es zweckmäßig, entsprechende Identifikationsinformationen des zugehörigen Teilnehmers bzw. der zugehörigen Set-Top-Box zusammen mit den Sprachdaten über den Rück- bzw. Sprachkanal zu senden.

Auf diese Weise kommt der Fachmann ausgehend von Druckschrift **NK11** in naheliegender Weise zum fünften Teilschritt des Patentanspruchs 2.

4.2.2 Der **NK11** ist ferner ein Ausführungsbeispiel zu entnehmen, bei dem der Teilnehmer im Rahmen eines elektronischen Einkaufsdienstes eine Lieferadresse („shipping address“) in „nicht-textueller“ Form über ein Grafiktablett eingibt. Zumindest wenn der Teilnehmer für sich selbst einkauft, stimmt die Lieferadresse mit der Adresse des Teilnehmers und dem Standort seiner Set-Top-Box überein, so dass sie einen Teilnehmerstandort im Sinne von Patentanspruch 1 bezeichnet. Der Server 13 konvertiert die Adresse in einen Text und sendet diesen an die auf der Set-Top-Box ausgeführte Einkaufsdienstanwendung, um ihn an dem Teilnehmerstandort anzuzeigen (**NK11**, Seite 10, Zeile 14 bis 19 i. V. m. Seite 1, Zeile 38 bis Seite 2, Zeile 5, Spalte 2, Zeile 33 bis 34 sowie Claims 1, 2 und 7 bis 10).

Als Alternative zur Informationseingabe mittels des Grafiktablets lehrt die **NK11**, dass der Teilnehmer die in Textdaten zu konvertierenden Informationen auch in ein Mikrofon sprechen kann (vgl. Seite 10, Zeile 20 bis 27 i. V. m. Claims 1 und 11). Das bedeutet, dass der Teilnehmer der Einkaufsdienstanwendung seine Adresse auch in gesprochener Form zuführen kann.

In diesem Fall wird der Teilnehmer die Textdaten, die ihm auf die netzwerkseitige Sprachverarbeitung hin angezeigt werden, auf Genauigkeit und Richtigkeit überprüfen (vgl. **NK11**, Seite 10, Zeile 26 und 27 - „The textual data is returned to the set-top box, where it can be displayed to the user. The user can [...] confirm that the text has been accurately generated from the voice data“). Im Zuge dieser Überprüfung muss er gedanklich die Information ermitteln, die die Textdaten

repräsentieren. Diese Information ist sein eigener Standort, falls der angezeigte Text seine korrekte Adresse darstellt.

Zudem ist es selbstverständlich, dass auch die elektronische Einkaufsdienstanwendung die von dem Server 13 in Textform übermittelte Lieferadresse auf Plausibilität überprüft (etwa um unvollständig erkannte Adressinformationen von einer Weiterverarbeitung auszuschließen) und erst danach die tatsächliche, „endgültige“ Lieferadresse - und damit den Standort des Teilnehmers - festlegt. Denn nur auf diese Weise ist gewährleistet, dass die von dem Teilnehmer erworbenen Produkte in jedem Fall an den korrekten Bestimmungsort geliefert werden (vgl. auch **NK11**, Seite 9, Zeile 27 bis 29 - „convert the entire image to text and then parse the text to determine the recipient's address“ für den Fall der Weiterverarbeitung der erkannten Texte zur Festlegung einer E-Mail-Adresse).

Alles in allem kann somit ausgehend von **NK11** auch ein Naheliegen des zweiten Teilschritts des Patentanspruchs 2 begründet werden.

4.2.3 Die Beklagte ist der Auffassung, gemäß **NK11** basiere die Erkennung der Versandadresse nicht auf einer netzwerkseitigen Erkennung eines Sprachinhalts, sondern lediglich auf einer Erkennung der Textzeichen, aus denen die Adresse besteht. Wie die Versandadresse erkannt werde, sei nicht offenbart; jedenfalls solle dies in jedem Fall auf der Set-Top-Box erfolgen.

Diese Argumentation greift nicht durch.

Denn mit Patentanspruch 2 („determining said associated user site [...]“) ist keine netzwerkseitige Erkennung einer Versandadresse, sondern eine Bestimmung, Festlegung oder Ermittlung des Teilnehmerstandorts beansprucht, die an einem beliebigen Ort vorgenommen werden kann und gemäß dem zweiten Teilschritt von Patentanspruch 2 auf den erkannten Sprachinhalten beruht („from said recognized speech content“). Dies ist - wie oben in Abschnitt **III.4.2.2** ausgeführt - aus der Lehre

der **NK11** abzuleiten. Zu dem Aspekt der netzwerkseitigen Erkennung von Sprachinhalten wird im Übrigen auf die Abschnitte **III.1.2** und **III.1.3** (s.o.) verwiesen.

4.3 Die vorstehenden Überlegungen aus Abschnitt **III.4.2** gelten entsprechend für die Merkmale von Patentanspruch 14, die somit ebenfalls keine erfinderische Tätigkeit begründen können.

5. Somit haben sämtliche angegriffenen Patentansprüche des Streitpatents keinen Bestand.

IV.

Das Streitpatent ist auch in keiner der Fassungen der Hilfsanträge 1 bis 8 bestandsfähig.

1. Der Gegenstand des Patentanspruchs 1 in der Fassung nach Hilfsantrag 1 beruht nicht auf erfinderischer Tätigkeit.

1.1 Patentanspruch 1 nach Hilfsantrag 1 unterscheidet sich von dem erteilten Patentanspruch 1 dadurch, dass auf Merkmal **1.2.5** noch das Merkmal

1.2.6 wherein the entertainment and information services comprise a Video-On-Demand service and an Interactive Program Guide.

folgt.

Dieses Merkmal führt mit einem Video-on-Demand-Dienst und einem interaktiven Programmführer spezielle Unterhaltungs- und Informationsdienste an, deren Lieferung gemäß Merkmal **1.2.5** individuell kontrolliert werden soll. Ein Video-on-Demand-Dienst ist aus fachmännischer Sicht eine Funktionalität eines Kommunikationsnetzwerks, die es ermöglicht, Videos auf Teilnehmeranfragen hin bereitzustellen.

1.2 Der Gegenstand von Patentanspruch 1 nach Hilfsantrag 1 beruht ausgehend von Druckschrift **NK12** auf keiner erfinderischen Tätigkeit.

1.2.1 Die **NK12** wurde im August 1995 - also vor dem ältesten Prioritätszeitpunkt - veröffentlicht und bildet daher Stand der Technik.

1.2.2 In **NK12** sind Methoden zur Verwendung von Spracherkennung beschrieben, um Fernsehbilder, Audiodaten und sonstige Daten auszuwählen oder zu verändern, die über ein Kabelfernsehsystem an einen Teilnehmer übertragen werden (vgl. den Druckschriftentitel „Speech Recognition Methods for Controlling Cable Television“ sowie Seite 285, erster Satz unter Figur 1). Um auf die zu übertragenden Informationen Einfluss zu nehmen, wird gemäß einer ersten Methode ein Telefonnetzwerk (Seite 285, Figuren 1 und 2 sowie Seite 285, zweiter Absatz unter Figur 2 bis Seite 287, zweiter vollständiger Absatz) und gemäß einer zweiten Methode das Kabelfernsehsystem selbst verwendet (Figur 3 mit Seite 287, vorletzter und letzter Absatz).

Bei der zweiten Methode wird in einem Haus des Teilnehmers (vgl. den gestrichelten Kasten in Figur 3 sowie Figur 1 i. V. m. Seite 285, letzter Absatz - „A user's house 1“) ein Mikrofon 24 mit einer Kabelbox 25 verbunden, von der aus akustische Daten bzw. Sprachbefehle eines Teilnehmers vermischt mit Daten anderer Teilnehmer auf einem Kanal oder Subkanal des Kabelfernsehsystems zum Kabelfernsehsystem 30 - also auf einem Rückkanal des Netzwerks - übertragen werden („From the cable box 25, acoustic data from the user is sent over the cable TV network 26, having been decoded partially or entirely [...] the user's voice commands are passed unaltered onto a channel or sub-channel of the cable television system, [...], are intermixed with transmissions from other users [...]).“).

Es liegt im Rahmen des fachmännischen Wissens, dass solche vermischten Teilnehmerdaten auf einem Rückkanal eines Kabelfernsehsystems üblicherweise mit Hilfe von Zeitmultiplexverfahren (z.B. TDMA-Verfahren) übertragen werden, bei denen Sprachdatenpakete verschiedener Teilnehmer in verschiedenen

Zeitschlitzten auf demselben Rückkanal gebündelt gesendet werden (s.o., Ausführungen in Abschnitt **III.1.1**).

Der Fachmann wird daher erkennen, dass das Vermischen der akustischen Daten bzw. Sprachbefehle des Teilnehmers mit den Daten anderer Teilnehmer insbesondere in der Anwendung eines TDMA-Zeitmultiplexverfahrens auf Sprachdatenpakete liegt, die von den anderen Teilnehmern stammen.

Somit beschreibt die Druckschrift **NK12** nach fachmännischem Verständnis ein Verfahren zur Verwendung eines Rückkanals des in Figur 3 dargestellten Kabelfernsehnetzwerks (Merkmal **1.1**). Ein solches Netzwerk umfasst in der Regel mehrere hundert oder tausend Teilnehmerstandorte (vgl. **NK9**, Spalte 22, Zeile 40 bis 43), so dass davon ausgegangen werden kann, dass sehr viele Teilnehmer die in **NK12** beschriebene Spracherkennungsfunktionalität in Anspruch nehmen.

Um die akustischen Daten bzw. die Sprachbefehle der Teilnehmer mittels eines TDMA-Verfahrens auf dem Rückkanal vermischt zu übertragen (und im Zuge dessen netzwerkseitig zu empfangen), ist es erforderlich, die einzelnen Sprachdatenpakete vorher festzulegen.

Somit sind auch die Merkmale **1.1.1**, **1.2** und **1.2.1** aus **NK12** ableitbar.

Die vermischten Sprachdatenpakete müssen vor der Verarbeitung im Spracherkennungssystem 28 wieder „entmischt“, d.h. aufgeteilt oder zerlegt werden, so dass die ursprünglichen „unvermischten“ und zusammenhängenden Sprachinformationen eines jeden Teilnehmers rekonstruiert werden können. Zu diesem Zweck sind bei Zeitmultiplexverfahren entsprechende Demultiplex-Schritte vorgesehen.

Die Notwendigkeit einer Aufteilung der vermischten Sprachdaten der Teilnehmer ergibt sich auch bereits daraus, dass eine Spracherkennungssoftware zeitlich geordnete Sprachdatenfolgen der einzelnen Teilnehmer benötigt und keine

sinnvollen Ergebnisse liefert, wenn sie auf zeitlich ineinander verschränkte („vermischte“) Sprachdatenpakete verschiedener Teilnehmer angewendet wird.

Somit liegt in der Lehre der **NK12** auch Merkmal **1.2.2** vor.

Des Weiteren lehrt Druckschrift **NK12**, dass die Sprachbefehle eines Teilnehmers an ein Spracherkennungssystem 28 geleitet werden (vgl. Seite 287, vorletzter Absatz i. V. m. Figur 3 - „voice commands are directed [...] to a speech recognition system 28“). Es ist selbstverständlich, dass das Sprachverarbeitungssystem 28 die übertragenen Sprachbefehle - zumindest zum Teil - auch tatsächlich verarbeitet, da es andernfalls seiner Funktion, Fernsehbilder, Audiodaten und sonstige Daten im Sinne des Teilnehmers auszuwählen und zu verändern (s.o.) nicht nachkäme und zudem keine Notwendigkeit bestünde, überhaupt ein Sprachverarbeitungssystem 28 vorzusehen.

Somit entnimmt der Fachmann der Druckschrift **NK12** auch das Merkmal **1.2.3**.

Aus **NK12** geht ferner hervor, dass das Spracherkennungssystem 28 den Fernsehsignalgenerator 29 und einen Teil des Kabelfernsehsystems 30 steuert, nachdem es die Sprachbefehle des Teilnehmers erhalten und erkannt hat (vgl. Seite 287, vorletzter Absatz - „voice commands are directed [...] to a speech recognition system 28, which in turn controls a television signal generator 29 and a portion of cable system 30“; vgl. auch den Titel „Speech Recognition Methods for Controlling Cable Television“ der **NK12**). Diese Befehle können als Steuerungsanfrage des Teilnehmers angesehen werden, die selbstverständlich an die zu steuernden Einheiten (Fernsehsignalgenerator 29, Teil des Kabelfernsehsystems 30) weitergegeben werden muss, so dass diese in Übereinstimmung mit den erkannten Sprachbefehlen gesteuert werden.

Auf diesem Wege führt die Druckschrift **NK12** den Fachmann auch zu Merkmal **1.2.4**.

Die Bereitstellung und Übertragung der drei Datenarten „Fernsehbilder“, „Audiodaten“ und „sonstige Daten“ kann jeweils als eine eigene Netzwerkfunktionalität und damit jeweils als ein eigener „Informationsdienst“ im Sinne des Merkmals **1.2.5** angesehen werden, wobei davon auszugehen ist, dass die Übertragung von Fernsehbildern oder Audiodaten zumindest zeitweise auch der Unterhaltung der Teilnehmer dient.

Somit lehrt die Druckschrift **NK12** die sprachgesteuerte Beeinflussung der Lieferung von Unterhaltungs- und Informationsdiensten auf Basis der an das Spracherkennungssystem 28 geleiteten und von diesem erkannten Sprachbefehle (Merkmal **1.2.5**).

Druckschrift **NK12** ist ferner zu entnehmen, dass das Spracherkennungssystem auf der der Kabelbox des Teilnehmers „gegenüberliegenden“ Seite des Kabelfernsehnetzwerks angeordnet ist, auf der sich auch der Fernsehsignalgenerator 29 und das Kabelfernsehsystem 30 befinden (vgl. Figur 3). Denn die Spracherkennung im Spracherkennungssystem 28 wird erst dann vorgenommen, wenn die akustischen Daten des Teilnehmers über das Kabelfernsehnetzwerk übertragen worden sind (vgl. Seite 287, vorletzter Absatz - „the user’s voice commands are passed unaltered onto a channel or sub-channel of the cable television system [...] In general, voice commands are directed [...] to a speech recognition system 28 [...]“). Daraus folgt, dass sich das Spracherkennungssystem, in das die Sprachbefehle der Teilnehmer eingegeben werden, an einem Knoten des Kabelfernsehnetzwerks befindet (**Teilmerkmal** von Merkmal **1.1.2**).

Der Druckschrift **NK12** ist allerdings nicht ausdrücklich zu entnehmen, dass dieser Knoten ein Leitungsknoten des Kabelfernsehnetzwerks ist (**verbleibendes Teilmerkmal** von Merkmal **1.1.2**).

Auch Merkmal **1.2.6** geht nicht unmittelbar aus **NK12** hervor.

1.2.3 Ausgehend von Druckschrift **NK12** gelangt der Fachmann zum Gegenstand von Patentanspruch 1 nach Hilfsantrag 1, ohne dabei erfinderisch tätig zu werden.

1.2.3.1 So ist der Fachmann immer bestrebt, bei der praktischen Einrichtung eines Kommunikationsnetzwerks Ressourcen zu sparen und gleichzeitig einen möglichst schnellen Datenaustausch zwischen den Komponenten des Netzwerks zu gewährleisten. Ihm ist bewusst, dass dies bei dem in **NK12** beschriebenen Kabelfernsehsystem auf einfache Weise dadurch zu erreichen ist, dass die in Figur 3 dargestellten netzwerkseitigen Komponenten 27 bis 30 in unmittelbarer Nähe zueinander untergebracht werden, da in diesem Fall nur kurze Übertragungsleitungen erforderlich sind und darüber hinaus eine gemeinsame Infrastruktur zur Verbindung, Unterbringung und Verwaltung dieser Komponenten genutzt werden kann.

Daher wird der Fachmann bei der technischen Umsetzung der aus Druckschrift **NK12** bekannten Lehre die Komponenten 27 bis 30 nahe zueinander anordnen (z.B. in demselben Gebäude). Nachdem die Komponenten 29 und 30 nach fachmännischem Verständnis als ein Headend des Kabelfernsehsystems angesehen werden können, das mit den Kabelboxen der Teilnehmer kabelgebunden kommuniziert, folgt daraus, dass sich das Spracherkennungssystem 28 in diesem Fall an einem Leitungsknoten des Kabelfernsehsystems befindet.

Somit gelangt der Fachmann auf naheliegende Weise zum **verbleibenden Teilmerkmal** von Merkmal **1.1.2**.

1.2.3.2 Weiterhin überlässt es die Druckschrift **NK12** dem Fachmann, die Fernsehbilder und die sonstigen Daten festzulegen, deren Lieferung durch die an die Komponenten 29 und 30 übermittelte Steuerungsanfrage des Teilnehmers beeinflusst wird.

Für den Fachmann liegt es auf der Hand, dass diese Daten insbesondere digitale Videos und elektronische Programmführer sein können (vgl. Spalte 5, Zeile 39

bis 53; Spalte 8, Zeile 44 bis 55 i. V. m. Spalte 8, Zeile 60 bis Spalte 11, Zeile 3 der Druckschrift **NK9**, in der diejenigen Informationen explizit aufgezählt werden, die gewöhnlich von einem Headend an einen Teilnehmer geliefert werden - „The information may include [...] digital video, [...], electronic program guides, [...]“).

Der Fachmann wird daher bei Bedarf die aus dem Fernsehsignalgenerator 29 und dem Kabelfernsehsystem 30 bestehende Netzwerkkomponente derart auslegen, dass ein Teilnehmer nicht nur die Übermittlung von Videos mittels seiner Sprachbefehle beeinflussen kann, sondern auch die Übermittlung von elektronischen Programmführerdaten - beispielsweise, um diejenigen EPG-Daten auszuwählen, die konkret geliefert werden sollen (vgl. **NK9**, Spalte 22, Zeile 43 bis 51 - „head-end installation 125 is preferably provided with the capability of selecting only certain portions of the EPG data to be transmitted to the subscriber terminal units“).

Der Fachmann gelangt daher auch zu Merkmal **1.2.6** auf naheliegende Weise.

1.3 Der Argumentation der Beklagten, der Gegenstand des Patentanspruchs 1 nach Hilfsantrag 1 beruhe auf einer erfinderischen Tätigkeit, kann nicht beigetreten werden.

1.3.1 Die Beklagte bringt sinngemäß vor, in **NK12** sei nicht offenbart, dass ein Rückkanal akustische Nutzersignale von verschiedenen Nutzern oder unterschiedlichen Standorten aufweist; so könnten sich mehrere Nutzer auch ein Mikrofon und einen Fernseher teilen. Auch wenn das Kabelfernsehsystem mehrere Teilnehmerstandorte umfasse, bedeute dies noch nicht, dass auch eine Sprachverarbeitung für mehrere Nutzer vorgesehen sei. Zudem werde in **NK12** nur eine lokale Sprachsignalverarbeitung vorgenommen, was dadurch gestützt werde, dass der Begriff „decoding“ im vorletzten Absatz auf Seite 287 (siehe „decoded partially or entirely in the cable box 25 or in the cable gateway“) als „sprachsignalverarbeitet“ zu übersetzen sei. Ferner sei in Figur 3 kein Switch (der mehrere Quellensignale von mehreren Nutzern auf einen einzigen Empfänger aufteile und somit Merkmal **1.2.2** realisieren könne) vorgesehen, sondern nur ein

Splitter (der ein Signal von einer Quelle auf mehrere Empfänger aufteile bzw. ein Signalteil eines Signals von einem anderen abspalte). Dies lasse nur den Schluss zu, dass ein externes Vermischen einer Mehrzahl von Sprachkanälen außerhalb des in Figur 3 gezeigten Teilnehmerstandorts nicht angedacht gewesen sei.

Da in Figur 3 der **NK12** keine Verbindung zwischen dem Fernsehsignalgenerator 29 und dem Kabelfernsehsystem 30 eingezeichnet und das System 29 nur über das Spracherkennungssystem 28 an den Splitter 27 angeschlossen sei, sei unklar, wie die Steuerung bezüglich des Systems 29 auszuführen ist. Auch werde gemäß **NK12** nur ein Teil des Systems 30 („a portion of the cable system 30“, vgl. Seite 287, vorletzter Absatz) gesteuert. Eine Kontrolle der Lieferung von Unterhaltungs- und Informationsdiensten würde eine - in **NK12** nicht gezeigte - Kontrolle bis zum Fernsehgerät bedingen und könne daher nicht beschrieben sein. Ein Verarbeiten der Sprachkanäle und ein Erzeugen einer eindeutigen Antwort gemäß den Merkmalen **1.2.3** und **1.2.4** sei ebenfalls nicht offenbart.

Die **NK12** schweige zudem darüber, welche Daten die auf dem Kabelfernsehsystem modifizierten „Bilder, Klang und Daten“ genau sein sollten. Das Auswählen und Modifizieren einzelner Bilder oder Audiodaten sei kein anspruchsgemäßer Unterhaltungs- und Informationsservice. Auch werde in **NK12** nicht in Erwägung gezogen, dass individuelle Information für jeden Nutzer bereitgestellt werde; vielmehr klammere die **NK12** gemäß dem letzten Absatz auf Seite 286 (in diesem geht es um die Bereitstellung von „shared video and private audio“) die individuelle Beschaffung von Videodaten für die erste Methode explizit aus.

Auch sei es nicht nahegelegt, die Sprachkanäle an einem anspruchsgemäßen Leitungsknoten einzugeben. Ob der Fachmann Ressourcen sparen wolle, sei der **NK12** nicht zu entnehmen. Eine Veranlassung hierfür fehle. Selbst wenn diese bestanden hätte, hätte der Fachmann eine nicht netzwerkbasierte Lösung erwogen (z.B. eine direkte Verbindung der Einheiten 28 und 30), nicht aber die anspruchsgemäße Lösung.

1.3.2 Die Ausführungen der Beklagten erweisen sich als nicht überzeugend.

So ist bei einem Kabelfernsehsystem gewöhnlich davon auszugehen, dass sehr viele Teilnehmer dessen Benutzerschnittstellenfunktionalitäten - wie etwa eine Spracherkennung und Sprachsteuerung - gleichzeitig in Anspruch nehmen (vgl. **NK9**, Spalte 22, Zeile 40 bis 43).

Dass die Spracherkennung gemäß **NK12** ganz oder teilweise netzwerkseitig vorgenommen werden kann, folgt bereits aus der Angabe „In general, voice commands are directed [...] to a speech recognition system 28“ auf Seite 287 und dem Umstand, dass ein Spracherkennungssystem diejenigen Daten, die es erhält, auch verarbeitet. Für eine netzwerkseitige Spracherkennung spricht im Übrigen auch die Angabe „the user's voice commands are passed unaltered onto a channel or sub-channel of the cable television system“ auf Seite 287. Die weitere Angabe „acoustic data from the user is sent over the cable TV network 26, having been decoded partially or entirely“ in demselben Absatz lässt allenfalls den Schluss zu, dass die akustischen Daten des Teilnehmers zum Teil am Teilnehmerstandort verarbeitet worden sind, so dass sie auch noch netzwerkseitig weiterverarbeitet werden müssen. In diesem Fall führt die Formulierung „intermixed with transmissions from other users“ vor dem Hintergrund des durch **NK16** dokumentierten Fachwissens des Fachmanns unmittelbar darauf, dass mehrere Sprachkanäle verschiedener Teilnehmerstandorte vermischt werden (s.o., Abschnitt **IV.1.2.2**).

Des Weiteren ist ein Switch grundsätzlich dadurch charakterisiert, dass er mehrere Quellensignale auf mehrere Empfänger verteilt. Bereits aus diesem Grund lässt sich aus dem Fehlen eines Switches nicht zwangsläufig auf ein teilnehmerseitiges Vermischen der Sprachsignale schließen. Zudem muss das Rücksignal im Sinne des Patentanspruchs 1 netzwerkseitig aufgeteilt („demultiplext“) werden, wenn die Sprachsignale mehrerer Teilnehmer wie in Abschnitt 5.3.1 der **NK16** beschrieben gebündelt („multiplext“) auf dem Rundfunk-Rückkanal übertragen werden. In diesem Fall kommt es nicht darauf an, ob das Nutzersignal durch einen Splitter oder einen Switch geleitet wird. Spezielle Ausgestaltungen der Komponenten, die das

Rücksignal empfangen und aufteilen (vgl. Merkmale **1.2.1**, **1.2.2**), sind nicht beansprucht.

Im Übrigen wird der Fachmann dem Splitter 27 insbesondere auch die Funktion des „Upstream splitter“ gemäß dem DOCSIS-Standard zuschreiben, der den Rückkanal, der 5 - 42 MHz Bandbreite umfasst, in verschiedene Datenströme („Data“) aufteilt (vgl. **NK19** (Data-Over-Cable Service Interface Specifications), Figur 1-2 auf Seite 3). Im Hinblick auf Figur 3 der **NK12** bedeutet das, dass der Splitter 27 zumindest die Sprachdaten des Rückkanals (diese werden an das Spracherkennungssystem 28 geleitet) von anderen Daten trennt, die ebenfalls über den Rückkanal an das Kabelfernsehsystem 30 geleitet werden.

Des Weiteren können die Merkmale **1.2.4** und **1.2.5** bereits aus dem Umstand abgeleitet werden, dass die Steuerung des Fernsehsignalgenerators 29 und des Teils des Kabelfernsehsystems 30 in Übereinstimmung mit den von dem Spracherkennungssystem 28 erkannten Sprachbefehlen vorgenommen wird (s.o., Abschnitt **IV.1.2.2**). Eine detailliertere Beschreibung des Zusammenwirkens der Komponenten 27 bis 30 ist hierfür nicht nötig.

Insoweit die **NK12** die Bereitstellung von „shared video and private audio“ lehrt, betrifft dies lediglich die erste aus **NK12** bekannte Methode. Abgesehen davon kann auch die Steuerung der Bereitstellung eines einzigen Videos für mehrere Teilnehmer („shared video“) als individuelle Kontrolle der Lieferung eines Video-on-Demand-Dienstes im Sinne von Patentanspruch 1 nach Hilfsantrag 1 angesehen werden, wenn diese von einem einzigen Teilnehmer - also „individuell“ - angestoßen wird; zu beachten ist in diesem Zusammenhang, dass Merkmal **1.2.5** nicht auf eine „individuelle Lieferung [...] zu einem einzigen Teilnehmerstandort“, sondern auf eine „individuelle Kontrolle der Lieferung [...] zu jedem Teilnehmerstandort“ abstellt.

Im Übrigen ergibt sich die Lehre der Merkmale **1.2.4** und **1.2.5** auch aus Druckschrift **NK9**, da jeder Teilnehmer sein Fernsehgerät sprachgesteuert ein- und ausschalten (Spalte 19, Zeile 61 bis Spalte 20, Zeile 9), Fernsehkanäle einstellen (Spalte 20, Zeile 62 bis Spalte 21, Zeile 6) oder durch die einzelnen Kanäle „surfen“ (Spalte 21,

Zeile 7 bis Spalte 22, Zeile 4) kann. Dabei wird als Reaktion auf die erkannten Sprachinhalte eines jeden Teilnehmers der Empfang von Fernsehprogrammen durch das Fernsehgerät des Teilnehmers - und damit die Lieferung der Programme an den Teilnehmerstandort im Sinne von Merkmal **1.2.5** - beeinflusst. Die Fernsehprogramme dienen selbstverständlich der Unterhaltung und Information der Teilnehmer.

Wie in Abschnitt **IV.1.2.3.1** ausgeführt, hatte der Fachmann auch Veranlassung, die Sprachkanäle an einem Leitungsknoten einzugeben. Dabei kommt es nur auf die relative räumliche Nähe der Einheiten 28 und 30 an; ob diese „direkt“ verbunden oder Teil einer Netzwerkinfrastruktur sind, ist in diesem Zusammenhang nicht relevant.

1.4 Mit Rücksicht auf die Ausführungen zum erteilten Patentanspruch 1 beruht die Lehre des Patentanspruchs 1 gemäß Hilfsantrag 1 somit auf keiner erfinderischen Tätigkeit und ist daher nicht patentfähig. Mit seinem Patentanspruch 1 fällt der gesamte Hilfsantrag 1.

Beantragt der Patentinhaber, das Patent in beschränktem Umfang mit einem bestimmten Anspruchssatz oder bestimmten Anspruchssätzen aufrechtzuerhalten, rechtfertigt es grundsätzlich die Ablehnung des gesamten Antrags, wenn sich auch nur der Gegenstand eines Patentanspruchs aus dem vom Patentinhaber verteidigten Anspruchssatz als nicht patentfähig erweist (BGH, Beschluss vom 27. Juni 2007, X ZB 6/05, juris und GRUR 2007, 862 – *Informationsübermittlungsverfahren II*). Allerdings ist das Gericht gehalten, aufzuklären, in welchem Verhältnis die Hilfsanträge zu einem nicht ausdrücklich formulierten Petitum stehen sollen, einem formal vorrangigen Antrag nur teilweise zu entsprechen (BGH, Urteil vom 13. September 2016, X ZR 64/14, juris und GRUR 2017, 57 – *Datengenerator*).

Im vorliegenden Fall hat die Beklagte in der mündlichen Verhandlung erklärt, dass sie die Patentansprüche der Hilfsanträge als jeweils geschlossene Anspruchssätze ansieht, die jeweils insgesamt beansprucht werden. Dies schließt für den

Hilfsantrag 1 sowie für alle übrigen Hilfsanträge eine separate Betrachtung einzelner Patentansprüche aus, wenn sich ein Patentanspruch des betroffenen Anspruchssatzes, wie hier, als nicht patentfähig erweist.

2. Auch Hilfsantrag 2 hat keinen Erfolg.

2.1 Die Verteidigung von Patentanspruch 1 nach Hilfsantrag 2 ist unzulässig.

Patentanspruch 1 nach Hilfsantrag 2 unterscheidet sich von dem erteilten Patentanspruch 1 dadurch, dass nach Merkmal **1.2.5** das Merkmal

1.2.7 wherein the speech recognition is performed only by said speech recognition system.

angehängt ist.

2.1.1 Nach der Rechtsprechung des Bundesgerichtshofs muss die Selbstbeschränkung des Patentinhabers im Nichtigkeitsverfahren immer zu einer (zulässigen) Einschränkung des Patents führen. Die Beschränkung ist auch durch Aufnahme eines Merkmals, das sich in Selbstverständlichkeiten erschöpft, nicht möglich, vielmehr muss dem Merkmal ein den Patentanspruch kennzeichnender, unterscheidungskräftiger Sinn zukommen. Änderungen, die keine Beschränkung bedeuten, sind grundsätzlich unzulässig (BGH, Urteil vom 14. September 2004, X ZR 149/01, juris und GRUR 2005, 145 – *elektronisches Modul*, I.2; BGH, Urteil vom 7. Juni 2006, X ZR 105/04, juris und GRUR 2006, 923 – *Luftabscheider für Milchsammelanlage*, Rn. 16; *Keukenschrijver*, a. a. O., Rn. 296, m. w. N.).

2.1.2 Bei Anwendung dieser Grundsätze ist die Aufnahme von Merkmal **1.2.7** in den Patentanspruch 1 des Streitpatents als unzulässig zu beurteilen.

Merkmal **1.2.7** bezieht sich auf das in Merkmal **1.1.2** genannte Spracherkennungssystem („said speech recognition system“), das sich an einem

Leitungsknoten des Netzwerks befindet und die in Merkmal **1.2.3** beschriebene Spracherkennung durchführt. Durch die Aufnahme von Merkmal **1.2.7** in den erteilten Patentanspruch 1 und insbesondere durch die Verwendung des bestimmten Artikels in der Angabe „the speech recognition“ wird zum Ausdruck gebracht, dass genau diejenigen Spracherkennungsoperationen, die gemäß Merkmal **1.2.3** ausgeführt werden sollen und die deshalb in Merkmal **1.2.3** angesprochen sind, nur von dem an dem Leitungsknoten angeordneten Spracherkennungssystem ausgeführt werden sollen. Von anderen Spracherkennungsoperationen ist in Patentanspruch 1 nach Hilfsantrag 2 nicht die Rede.

Bereits der erteilte Patentanspruch 1 - ohne Merkmal **1.2.7** - ist jedoch so zu verstehen, dass die in Merkmal **1.2.3** genannten Spracherkennungsoperationen ausschließlich von dem in Merkmal **1.1.2** genannten Spracherkennungssystem an dem Leitungsknoten ausgeführt werden (s.o., Abschnitt **II.4.5**).

Somit wird der Gegenstand des erteilten Patentanspruchs 1 durch die Aufnahme von Merkmal **1.2.7** nicht eingeschränkt, so dass dem Merkmal **1.2.7** kein hinreichend unterscheidungskräftiger Sinn zukommt.

2.2 Unabhängig davon beruht die Lehre des Patentanspruchs 1 gemäß Hilfsantrag 2 nicht auf erfinderischer Tätigkeit.

So geht Merkmal **1.2.7** aus **NK11** hervor (vgl. Figur 1 i. V. m. Seite 10, Zeile 25 und 26 sowie Seite 2, Zeile 15 bis 25).

Unter Berücksichtigung der Ausführungen zum Hauptantrag ist die Lehre des Patentanspruchs 1 nach Hilfsantrag 2 somit ausgehend von **NK11** nahegelegt.

3. Hilfsantrag 3 kann nicht günstiger beurteilt werden, da der Gegenstand seines Patentanspruchs 1 ausgehend von Druckschrift **NK11** in Verbindung mit dem aus Druckschrift **NK9** entnehmbaren Stand der Technik nahegelegt ist.

Patentanspruch 1 nach Hilfsantrag 3 unterscheidet sich von dem erteilten Patentanspruch 1 dadurch, dass auf Merkmal **1.2.5** die Merkmale

1.2.8 wherein each user site comprises at least one set top box which is associated with a remote control comprising a microphone and a talk button;

1.2.9 and when the talk button is pushed by the user, the remote control sends a talk-button-active command to the set top box;

und

1.2.10 and the set-top box then places an icon on a screen coupled to the set top box and/or otherwise indicates to the user that she or he is listened to.

folgen.

Merkmal **1.2.8** besagt, dass jeder Teilnehmerstandort mindestens eine Set-Top-Box umfasst, die mit einer Fernbedienung verbunden ist oder zu einer Fernbedienung gehört („is associated with a remote control“), welche ein Mikrofon und eine Sprechtaaste („talk button“) aufweist. Die Sprechtaaste ist in der Streitpatentschrift auch als „push-to-talk button“ bezeichnet (vgl. Absätze [0039], [0059], [0060], [0089], [0090], [0095]).

Aus fachmännischer Sicht ist eine Set-Top-Box ein Gerät, das mit einem anderen Gerät verbunden ist und einem Benutzer dadurch zusätzliche Nutzungsmöglichkeiten bietet. Eine Set-Top-Box ist üblicherweise mit einem Fernseher verbunden und wandelt das Signal einer externen Signalquelle in einen Inhalt um, der auf dem Fernseher dargestellt wird.

Mit Merkmal **1.2.9** wird beansprucht, dass die Fernbedienung einen „Sprechtaste aktiv“-Befehl (ein „talk-button-active command“) an die Set-Top-Box sendet, wenn der Teilnehmer die Sprechtaste drückt. Ein solcher Befehl dürfte dem in der Beschreibung erwähnten „PTT active“-Befehl entsprechen, der allerdings an ein Set-Top-Gerät („set top appliance“) gesendet wird, das sich in der Nähe der Set-Top-Box befindet (vgl. Absatz [0090] i. V. m. Absatz [0089]).

Das Streitpatent lässt offen, welche Aktion durch den „Sprechtaste aktiv“-Befehl konkret angewiesen werden soll, führt jedoch verschiedene Vorgänge an, die nach dem Drücken der Sprechtaste ausgeführt werden. Dazu gehört das Anbringen einer digitalen Adresse am Anfang eines Sprachpakets, das Senden von Daten an eine Zentralstelle (Absatz [0039]), das Starten des Spracherkennungsprozesses, indem „das System“ darüber informiert wird, dass der Teilnehmer gleich sprechen wird (Absatz [0060]), oder - wie mit Merkmal **1.2.10** beansprucht - das Anordnen eines Icons auf einem mit der Set-Top-Box gekoppelten Bildschirm und/oder das Anzeigen, dass dem Teilnehmer zugehört wird (Absatz [0090]).

Demzufolge signalisiert der „Sprechtaste aktiv“-Befehl, dass die Sprechtaste gedrückt („aktiv“) ist; er löst insbesondere die in Merkmal **1.2.10** beschriebenen Aktionen aus.

3.1 Ausgehend von Druckschrift **NK11** gelangt der Fachmann in Verbindung mit der Lehre der Druckschrift **NK9** zum Gegenstand des Patentanspruchs 1 nach Hilfsantrag 3, ohne erfinderisch tätig zu werden.

3.1.1 Die Druckschrift **NK9** zeigt ein Abonnementfernsehsystem 100, bei dem Informationsprovider 114-n Programminhalte über einen Satelliten 123 und ein Headend 125 an Abonnentenstandorte 120-n liefern (vgl. Figuren 2A-C i. V. m. Spalte 8, Zeile 39 bis Spalte 12, Zeile 26). Jeder dieser Standorte umfasst eine Terminaleinheit 160 (vgl. Figuren 2C und 3), auf die ein charakteristisches Vokabular geladen wird, so dass die Terminaleinheit eine Spracherkennung ausführen kann, um ein mit der Terminaleinheit gekoppeltes Fernsehgerät sprachbasiert zu steuern oder auf einen elektronischen Programmführer

sprachbasiert zuzugreifen (**NK9**, Spalte 14, Zeile 61 bis Spalte 15, Zeile 18 i. V. m. Spalte 12, Zeile 5 bis 13; Spalte 15, Zeile 42 bis 46 und Zeile 55 bis 59; Spalte 19, Zeile 5 bis Spalte 32, Zeile 11, insbesondere Spalte 19, Zeile 22 bis 24, Spalte 22, Zeile 19 bis 21, Spalte 24, Zeile 49 bis 51 - „the user may speak sounds or words for controlling the EPG, [...] and television [...]“; „EPG“ = „electronic program guide“).

Gemäß **NK9** befindet sich an jedem Teilnehmerstandort 120-1 bis 120-n eine Terminaleinheit 160 mit einer zugehörigen Fernbedienung 166, die ein Mikrofon 320 und einen „<Recognize> button 772“ - d.h. eine Sprech- oder Spracherkennungstaste - aufweist (vgl. Figuren 2C, 4 und 9 i. V. m. Spalte 11, Zeile 65 bis Spalte 12, Zeile 13; Spalte 17, Zeile 17 bis 22 und Zeile 49 bis 53; Spalte 19, Zeile 5 bis 9).

Die Terminaleinheit 160 ist mit einem Fernsehgerät verbunden, empfängt ein Fernsehsignal, das von einem Headend übertragen wird, und gibt dieses an ein Fernsehgerät aus (Figur 3 i. V. m. Spalte 12, Zeile 5 bis 10 sowie Spalte 10, Zeile 65 bis Spalte 11, Zeile 6 sowie Spalte 12, Zeile 27 bis Spalte 13, Zeile 62). Somit ist die Terminaleinheit eine Set-Top-Box im Sinne des Streitpatents.

Der Spracherkennungsvorgang (die „speech recognition operation“) wird im Hauptprozessor 200 der Terminaleinheit aktiviert, indem der „<Recognize> button 772“ gedrückt wird (Spalte 19, Zeile 6 bis 9 i. V. m. Spalte 15, Zeile 42 bis 46; s. auch Spalte 33, Zeile 49 bis 53 sowie Spalte 19, Zeile 61 bis 64 i. V. m. Spalte 5, Zeile 1 bis 7). Somit kann der „<Recognize> button 772“ als Sprechstaste im Sinne von Merkmal **1.2.8** gelten.

Die Druckschrift **NK9** zeigt damit das Merkmal **1.2.8**.

Die Terminaleinheit 160 zeigt an, dass die Spracherkennung aktiviert ist, indem beispielsweise die Anzeige des Texts „Listening ...“ auf dem Bildschirm des Fernsehgeräts veranlasst wird (Spalte 19, Zeile 12 bis 21 - „[...] the recognition operation is activated [...] When activated, subscriber terminal unit 160 preferably

provides a suitable indication to this effect to the user [...] superimpose a display of the word „Listening ...“ [...]). Das bedeutet, dass die Information über die Aktivierung der Spracherkennung von der Fernbedienung an die Set-Top-Box übertragen werden muss, um anschließend die Anzeige des Texts auszulösen. Für den Fachmann ist es selbstverständlich, dass dazu ein entsprechender Befehl von der Fernbedienung an die Set-Top-Box gesendet wird - ebenso wie beim Wählen der Fernsehkanäle oder beim Einstellen der Lautstärke (vgl. **NK9**, Spalte 16, Zeile 51 bis 57).

Ein solcher Befehl kann als „Sprechtaste-aktiv“-Befehl im Sinne von Merkmal **1.2.9** bezeichnet werden, da er an die Set-Top-Box übertragen wird, wenn die Sprechtaste gedrückt worden ist.

Somit geht auch Merkmal **1.2.9** aus der **NK9** hervor.

Ferner vermittelt das Anzeigen des Textes „Listening ...“ dem Teilnehmer die Information, dass ihm „zugehört“ wird.

Damit ist auch die zweite Alternative von Merkmal **1.2.10** der **NK9** zu entnehmen.

3.1.2 Die Lehre der Patentanspruchs 1 gemäß Hilfsantrag 3 ist durch den aus den Druckschriften **NK11** und **NK9** entnehmbaren Stand der Technik nahegelegt.

So lehrt die Druckschrift **NK11**, dass ein spezielles Mikrofon in eine Fernbedienung eingebaut werden kann, so dass ein Teilnehmer Spracheingaben vornehmen kann (**NK11**, Seite 10, Zeile 20 bis 23), überlässt dem Fachmann jedoch die Einzelheiten der technischen Umsetzung dieser Eingabemöglichkeit. Hiervon ausgehend sieht sich der Fachmann im Stand der Technik nach Lösungen um, die die Implementierung einer Spracheingabe über ein Mikrofon einer Fernbedienung in einem Kabelfernsehnetzwerk näher beschreiben und stößt so auf die **NK9**. Diese zeigt verschiedene Möglichkeiten auf, wie Informationen, die von einem Teilnehmer eines Kabelfernsehnetzwerks in ein Mikrofon gesprochen worden sind, von einer

Spracherkennungsschnittstelle an einem Teilnehmerstandort weiterverarbeitet werden (vgl. Figuren 4 bis 6 i. V. m. Spalte 15, Zeile 19 bis Spalte 16, Zeile 50).

Dabei wird durch die Verwendung der Sprechaste 772 der Stromverbrauch der Fernbedienung gesenkt (Spalte 17, Zeile 16 bis 22); zudem wird die Benutzerfreundlichkeit verbessert, da dem Teilnehmer durch den angezeigten Text „Listening ...“ (s.o., Abschnitt **IV.3.1.1**) eine Rückmeldung darüber gegeben wird, dass seine mündlichen Äußerungen von der Fernbedienung und der Set-Top-Box korrekt erfasst werden.

Um sich diese Vorteile zunutze zu machen, wird der Fachmann daher die in **NK11** beschriebene Fernbedienung mit einer Sprechaste versehen und wie in **NK9** beschrieben mit der Set-Top-Box koppeln. Er gelangt somit in naheliegender Weise zu den Merkmalen **1.2.8** bis **1.2.10**.

3.1.3 Die Beklagte bringt vor, gemäß **NK9** würden zwar spracherkannte Wörter oder Befehle auf einem Bildschirm hervorgehoben, aber nur dann, wenn sie sowieso auf dem Bildschirm angezeigt würden. Dadurch sei jedoch noch keine Fernbedienungsfunktion gemäß Merkmal **1.2.9** nahegelegt. Zudem werde ein Symbol nur im Rahmen eines Ausführungsbeispiels angezeigt, bei dem ein Teilnehmer zur Aktivierung der Spracherkennung den Befehl „ATTENTION“ spreche und kein „<Recognize> button“ vorgesehen sei. Die Anzeige eines Symbols sei aber nur in diesem Fall sinnvoll, da der Sprecher ansonsten keine Rückmeldung erhalte, ob seine Worte erfasst werden. Werde hingegen ein „<Recognize> button“ verwendet, um die Spracherkennung zu aktivieren, sei es nicht erforderlich, ein Symbol anzuzeigen.

Diese Argumente greifen jedoch nicht durch.

So geht aus **NK9** unmittelbar hervor, dass die Spracherkennungsfunktion der Set-Top-Box entweder durch Drücken der Sprechaste 772 oder durch Sprechen des Befehls „ATTENTION“ aktiviert wird, was in beiden Fällen zur Anzeige des Texts „Listening ...“ führt (vgl. Spalte 19, Zeile 5 bis 14 - „when a <Recognize> button 772

is present [...], the recognition operation [...] is activated by pressing the <Recognize> button 772. When no <Recognize> button is present on the remote control 166, the recognition operation is activated by speaking an activation command such as „ATTENTION“. When activated, subscriber terminal unit 160 preferably provides a suitable indication to this effect to the user [...]).

Darüber hinaus ist es in jedem Fall auch sinnvoll, den Text „Listening ...“ anzuzeigen, wenn die Sprechaste 772 zur Aktivierung der Spracherkennung verwendet wird, da der Teilnehmer auf diese Weise eine Rückmeldung erhält, ob sein Tastendruck von der Fernbedienung und der Set-Top-Box tatsächlich erfasst worden ist.

3.2 Mit Rücksicht auf die Ausführungen zum erteilten Patentanspruch 1 ist der Gegenstand des Patentanspruchs 1 gemäß Hilfsantrag 3 somit nicht patentfähig. Mit dem Patentanspruch 1 fällt der gesamte Hilfsantrag 3.

4. Hilfsantrag 4 hat keinen Erfolg, da sich der Gegenstand seines Patentanspruchs 1 ebenfalls in naheliegender Weise aus einer Kombination der jeweiligen Lehren der Druckschriften **NK11** und **NK9** ergibt.

Patentanspruch 1 nach Hilfsantrag 4 ergibt sich aus dem erteilten Patentanspruch 1, indem hinter Merkmal **1.2.5** noch die beiden Merkmale

1.2.8' wherein each user site contains at least one set top box which is associated with a remote control containing a microphone and a talk button;

und

1.2.11 and wherein upon depressing the talk button on the remote control data are sent to the wireline node alerting the system as to the user site and a potential input.

angefügt werden.

Merkmal **1.2.8'** unterscheidet sich von Merkmal **1.2.8** dadurch, dass der Ausdruck „comprises“ durch den Ausdruck „contains“ ersetzt wird.

Der erste Teil von Merkmal **1.2.11** („upon depressing the talk button on the remote control data are sent to the wireline node“) bringt zum Ausdruck, dass Daten an den Leitungsknoten gesendet werden, wenn die Sprechaste gedrückt wird (vgl. auch Absatz [0039] der Streitpatentschrift - „when the button [...] is depressed, alerting the system [...])“).

Gemäß dem verbleibenden Teil von Merkmal **1.2.11** soll „das System“ in Übereinstimmung mit dem Sinngehalt des Begriffs „alerting“ im Hinblick auf den Standort und eine mögliche Eingabe des Teilnehmers benachrichtigt, informiert, gewarnt, alarmiert oder in Bereitschaft versetzt werden (vgl. auch Streitpatentschrift, Absatz [0060] - „[...] by informing the system that the subscriber is about to speak“). Unter diesem „System“ kann zum einen entsprechend dem Wortlaut von Patentanspruch 1 das Spracherkennungssystem zu verstehen sein, zum anderen kommt als „System“ auch das in Absatz [0034] der Streitpatentschrift genannte „multi-user control system for audio visual devices that incorporates a speech recognition system, [...] which may include a Cable Television (CATV) Headend“ in Frage, auf das jeweils in den Absätzen [0035] bis [0039] Bezug genommen wird.

Ein „alerting the system [...] as to the user site and a potential input“ ist somit insbesondere als ein Übertragen von Informationen an den Leitungsknoten (z.B. an ein Headend) zu verstehen, die den Teilnehmerstandort und eine Spracheingabe des Teilnehmers betreffen, die dieser möglicherweise vornehmen wird oder bereits vorgenommen hat, bevor die zugehörigen Sprachdaten am Spracherkennungssystem eingetroffen sind. Eine solche Information mit Bezug auf einen Teilnehmerstandort ist z.B. die Adresse einer Set-Top-Box (vgl. Streitpatentschrift, Absatz [0060], letzter Satz).

Merkmal **1.2.11** ist also insbesondere dann erfüllt, wenn der Leitungsknoten auf einen Sprechastendruck hin die Adresse einer Set-Top-Box erhält und ferner im

Hinblick auf eine möglicherweise folgende Spracheingabe eines Teilnehmers benachrichtigt wird.

4.1 Die Merkmale **1.2.8'** und **1.2.11** können eine Patentfähigkeit nicht begründen.

4.1.1 Merkmal **1.2.8'** geht inhaltlich nicht über Merkmal **1.2.8** hinaus und geht damit aus **NK9** hervor (vgl. Ausführungen zu Merkmal **1.2.8** in Abschnitt **IV.3.1.1**).

4.1.2 Merkmal **1.2.11** wird durch den aus den Druckschriften **NK11** und **NK9** bekannten Stand der Technik nahegelegt.

4.1.2.1 Das **erste Teilmerkmal** „wherein upon depressing the talk button on the remote control data are sent to the wireline node alerting the system as to [...] a potential input“ von Merkmal **1.2.11** beruht aus den folgenden Gründen auf keiner erfinderischen Tätigkeit:

Gemäß **NK11** wird die Spracherkennung an der Sendestation 10 von dem Server 13 vorgenommen, da zum Anmeldetag der **NK11** komplexe und rechenaufwendige Anwendungen nicht in Set-Top-Boxen implementiert werden konnten (Seite 10, Zeile 25 bis 26 i. V. m. Seite 2, Zeile 15 bis 25). Eine solche serverseitige Spracherkennung ist auch in Übereinstimmung mit der Lehre der **NK9** (vgl. Spalte 33, Zeile 49 bis 53 - “speech recognition operation [...] this processing could take place elsewhere in the system”).

Daher wird der Fachmann zumindest einen komplexen und rechenaufwendigen Spracherkennungsalgorithmus an der Sendestation des Kabelfernsehsystems ausführen. Das bedeutet, dass der Spracherkennungsvorgang (die „speech recognition operation“) auf die Betätigung der Sprech taste 772 hin nicht - wie in **NK9** ausführlich beschrieben - von dem Hauptprozessor 200 der Terminaleinheit 160, sondern von dem Server 13 der Sendestation 10 ausgeführt wird. Dazu muss die Terminaleinheit 160 die akustischen Daten, die von dem Mikrofon der

Fernbedienung aufgenommen werden, auf dem Rundfunk-Rückkanal des Kabelfernsehsystems zur Sendestation 10 übertragen.

a) Da ein Teilnehmer normalerweise nicht exakt zeitgleich mit dem Drücken der Sprechaste zu sprechen beginnt, enthalten die ersten Datenpakete mit akustischer Information, die nach dem Drücken der Sprechaste von der Terminaleinheit an die Sendestation übertragen werden, im Allgemeinen noch keine Sprachinformationen; diese Datenpakete stellen jedoch Informationen dar, die der Sendestation signalisieren, dass möglicherweise weitere Datenpakete mit Sprachinformationen folgen werden - nämlich dann, wenn der Teilnehmer die Sprechaste nicht gleich wieder loslässt, sondern auch tatsächlich in das Mikrofon spricht. Da das Übertragen der ersten Datenpakete ausgelöst wird, wenn die Sprechaste gedrückt worden ist, wird auf diese Weise das **erste Teilmerkmal** von Merkmal **1.2.11** verwirklicht.

b) Im Übrigen besteht die Notwendigkeit, eine Kommunikationsverbindung zwischen der Terminaleinheit und den Komponenten der Sendestation 10 zu eröffnen, die für den Empfang von Nachrichten von der Terminaleinheit vorgesehen sind, wenn der Teilnehmer die Sprechaste 772 drückt. Da nicht vorhersehbar ist, ob und wann ein Teilnehmer die Spracherkennungsfunktionalität seiner Set-Top-Box in Anspruch nimmt und dazu die Sprechaste betätigt, ist es zweckmäßig, eine Nachricht (einen „alert“) von der Terminaleinheit an die Komponenten der Sendestation zu senden, durch die die Eröffnung der Kommunikationsverbindung initiiert wird.

Für das Auslösen der Übersendung dieser Initiierungsnachricht bieten sich folgende Zeitpunkte an: wenn (i) die Sprechaste gedrückt wird, wenn (ii) der Teilnehmer bei gedrückter Sprechaste zu sprechen beginnt oder (iii) zu sprechen aufhört oder wenn (iv) die gedrückte Sprechaste wieder losgelassen wird.

ba) Die Alternative (i) hat gegenüber den anderen Alternativen den Vorteil, dass mit der Übertragung der Sprachdaten sobald wie möglich begonnen werden kann, so dass Verzögerungen, die durch die Eröffnung der Kommunikationsverbindung

verursacht werden, nicht oder nur in geringem Ausmaß ins Gewicht fallen und ferner Programmteile und/oder Daten zur Sprecheridentifikation (vgl. **NK9**, Spalte 19, Zeile 52 bis 55) frühzeitig von dem Spracherkennungssystem geladen werden können.

Für die Alternative (i) spricht zudem, dass aus Spalte 33, Zeile 49 bis 53 i. V. m. Spalte 19, Zeile 6 bis 9 und Spalte 15, Zeile 42 bis 46 der **NK9** ("speech recognition operation [...] this processing could take place elsewhere in the system"; "the recognition operation of the instant invention is activated by pressing the <Recognize> button 772"; "Main processor 200 performs the speech recognition operation") unmittelbar hervorgeht, dass die Spracherkennung an dem „anderen Ort“ („elsewhere“; hier: in der Sendestation 10) durch Drücken der Sprechaste aktiviert werden kann, wozu es ebenfalls erforderlich ist, auf die Betätigung der Sprechaste hin eine entsprechende Nachricht (einen „alert“) von der Fernbedienung an die Sendestation 10 zu übertragen.

Um sich die oben genannten Vorteile zunutze zu machen, wird der Fachmann die Übermittlung der Initiierungsnachricht zur Eröffnung der Kommunikationsverbindung insbesondere dann anstoßen, wenn der Teilnehmer die Sprechaste gedrückt hat.

bb) Jedoch ist nicht abzustreiten, dass auch die Alternativen (ii) bis (iv) spezifische Vorteile besitzen können. Wird die Übersendung der Initiierungsnachricht gemäß dieser weiteren Alternativen ausgelöst, reduziert sich nämlich gegenüber Alternative (i) die Zeitdauer, über die die Kommunikationsverbindung besteht, aber noch keine relevanten Sprachinformationen übertragen werden, was sich positiv auf die effektiv über das Netzwerk zur Verfügung stehende Datenübertragungskapazität auswirkt. Die Alternativen (ii) und (iii) haben allerdings den Nachteil, dass in der Set-Top-Box erkannt werden muss, ob der Teilnehmer gerade spricht; im Fall (iv) kann mit der netzwerkseitigen Spracherkennung erst sehr spät begonnen werden, was gerade bei umfangreicheren Äußerungen eines Teilnehmers den gesamten Sprachverarbeitungsvorgang unnötig verzögern kann.

Welche der Alternativen (i) bis (iv) besonders vorzugswürdig ist, hängt auch von dem Zeitaufwand ab, der zur Vorverarbeitung der Sprachsignale in der Set-Top-Box und zur Eröffnung der Kommunikationsverbindung zwischen der Set-Top-Box und den Komponenten der Sendestation aufgebracht werden muss.

Daher wird der Fachmann anhand der Umstände des Einzelfalls abwägen, welche Alternative zu wählen ist.

Die Wahl, die Initiierungsnachricht durch den Tastendruck des Teilnehmers auszulösen - und damit die Realisierung des **ersten Teilmerkmals** „wherein upon depressing the talk button on the remote control data are sent to the wireline node alerting the system as to [...] a potential input“ von Merkmal **1.2.11** - beruht damit aber allenfalls auf fachmännischen Abwägungen von bekannten Vor- und Nachteilen. Ein erfinderisches Zutun ist hierzu nicht erforderlich (vgl. BGH, Urteil vom 3. Mai 2006, X ZR 24/03, juris und GRUR 2006, 930 – *Mikrotom*, Leitsatz 1 und Rn. 33).

4.1.2.2 Dadurch, dass der Fachmann zusammen mit den Sprachdatenpaketen auch die MAC- und/oder NSAP-Adresse(n) der Set-Top-Box an die Sendestation überträgt (s.o., Abschnitt **III.4.2.1**), wird auch das **verbleibende Teilmerkmal** “upon depressing the talk button on the remote control data are sent to the wireline node alerting the system as to the user site” von Merkmal **1.2.11** verwirklicht.

4.2 Die Beklagte führt sinngemäß aus, die **NK9** zeige kein Senden eines direkten Steuersignals an einen Netzwerkknoten. Ausgehend von **NK9** habe der Fachmann weder Veranlassung, Merkmal **1.2.11** zu realisieren, noch seien ihm die Mittel an die Hand gegeben, dieses Merkmal umzusetzen. Zudem würde ein frühes Eröffnen einer Session unnötig Übertragungsbandbreite in Anspruch nehmen und Ressourcen verbrauchen; abgesehen davon stehe dem Fachmann eine überaus große Anzahl von Möglichkeiten zur Verfügung, die Datenübertragung zwischen der Set-Top-Box und der Sendestation 10 zu organisieren.

Ferner sei es technisch unsinnig, die der Analog-Digital-Wandlung dienenden sprachbezogenen Schaltkreise 322, 324 und 340, die die Sprachsignale für die Signalverarbeitung vorbereiteten (vgl. **NK9**, Figuren 5 und 6), netzwerkseitig zu implementieren, da ansonsten ein analoges Signal zum Knoten 517 übertragen werden müsste.

Die **NK9** stelle den Fachmann auch nicht vor die Aufgabe, Energie zu sparen. Eine „gepoolte“ Spracherkennung benötige keine Energiesparfunktion, da die Leerlaufzeiten einer Spracherkennung für mehrere Set-Top-Boxen im Vergleich zu denen einer Spracherkennung für eine einzelne Set-Top-Box vernachlässigbar seien. Wollte der Fachmann den Energieverbrauch netzwerkseitig optimieren, würde er allenfalls ganze Server zu- und abschalten.

Auch dieser Argumentation kann sich der Senat im Ergebnis nicht anschließen.

Es trifft zwar zu, dass die **NK9** keine detaillierten Aussagen zur Kommunikation zwischen einer Set-Top-Box und einem Netzknoten macht. Wie oben ausgeführt, hatte der Fachmann jedoch Veranlassung, Merkmal **1.2.11** zu realisieren, ohne dabei mit einer allzu großen Anzahl von Alternativen zur Auswahl des Auslösezeitpunkts der Initiierungsnachricht konfrontiert zu sein.

Dass der Fachmann bei seinen Überlegungen insbesondere den Zeitpunkt (i) in Betracht gezogen hätte und auch zur Umsetzung der entsprechenden Maßnahmen imstande gewesen wäre, wird im Übrigen durch die Aussage des Sachverständigen der Beklagten in Dokument **NK18e** bestätigt (dort Rn. 231 bis 233). Die zur technischen Umsetzung von Merkmal **1.2.11** erforderliche Gestaltung der Kommunikationsverbindungen zwischen den elektronischen Komponenten Fernbedienung, Set-Top-Box und Sendestation gehört zu den Routinetätigkeiten des Fachmanns.

Des Weiteren lässt der Umstand, dass die Spracherkennung in der Sendestation ausgeführt wird, nicht darauf schließen, dass auch die vorbereitenden Operationen der Schaltkreise 322, 324 und 340 netzwerkseitig ausgeführt werden müssten und

somit ein zwangsläufig ein analoges Signal zu übertragen wäre. Denn Merkmal **1.2.11** („upon depressing [...]“ = „beim Drücken“, „nach dem Drücken“, „wenn [...] gedrückt worden ist“) erfordert lediglich, dass das Senden der Daten in gewisser zeitlicher Nähe zum Tastendruck des Teilnehmers erfolgt. Patentanspruch 1 nach Hilfsantrag 4 umfasst also zumindest all diejenigen Fälle, in denen das analoge Sprachsignal nach der Betätigung der Sprechaste in dem Mikrofon oder in der Set-Top-Box umgehend in ein digitales Signal konvertiert wird (vgl. **NK9**, Figuren 4 und 5 mit Spalte 15, Zeile 24 bis 29 sowie Spalte 15, Zeile 64 bis Spalte 16, Zeile 6), so dass kein analoges Signal zur Sendestation übertragen werden muss.

Im Übrigen geht die Deaktivierung einer netzwerkseitigen Spracherkennungsfunktion, die einer Set-Top-Box zugeordnet ist, durch das Drücken der Sprechaste ausgelöst wird und die Übersendung einer entsprechenden Aktivierungsnachricht an die Sendestation erfordert, aus **NK9** unmittelbar hervor (s.o., Abschnitt **IV.4.1.2.1 ba**). Es erscheint nicht völlig ausgeschlossen, dass der Fachmann zumindest in bestimmten Spezialfällen (z.B. wenn nur ein einziger Spracherkennungsserver mit sehr vielen, jeweils einer Set-Top-Box zugeordneten parallelen Prozessoren zur Ausführung eines rechenaufwendigen Spracherkennungsverfahrens vorgesehen ist, wobei die Auslastung des Servers tagsüber sehr hoch, aber nachts sehr gering ist) eine teilnehmerabhängige, über das Kommunikationsnetzwerk vorgenommene Deaktivierung und Aktivierung einer Spracherkennungsfunktion - etwa durch Ab- und Zuschalten einzelner Prozessoren - auch tatsächlich realisiert hätte.

4.3 Mit Rücksicht auf die Ausführungen zum erteilten Patentanspruch 1 beruht die Lehre des Patentanspruchs 1 gemäß Hilfsantrag 4 somit auf keiner erfinderischen Tätigkeit und ist daher nicht patentfähig. Mit seinem Patentanspruch 1 fällt der gesamte Hilfsantrag 4.

5. Auch **Hilfsantrag 5** hat keinen Erfolg, weil der Gegenstand seines Patentanspruchs 1 durch die kombinierte Lehre der Druckschriften **NK11** und **NK9** nahegelegt ist.

Patentanspruch 1 nach Hilfsantrag 5 unterscheidet sich von Patentanspruch 1 nach Hilfsantrag 4 dadurch, dass vor dem Merkmal **1.2.8'** das Merkmal

1.2.12 wherein said network comprises a content engine from which said speech recognition system receives content status information;

eingefügt ist und zwischen den Merkmalen **1.2.8'** und **1.2.11** die beiden Merkmale

1.2.13 wherein the analog signals picked up by the microphone are pre-processed by the remote control;

und

1.2.14 wherein the set top box receives a radio frequency (RF) signal or an infra-red signal from the remote control; and

eingeschoben sind.

Laut Streitpatent sind die in Merkmal **1.2.12** genannten Inhaltsstatusinformationen („content status information“) beispielweise Listen zwischengespeicherter Inhaltsdaten (etwa von interaktiven Programmführern), Listen verfügbarer Dienste oder auch Videosequenzen, die durch Video-on-Demand- oder Pay-per-View-Dienste bereitgestellt werden (Streitpatentschrift, Absätze [0268], [0269], [0316], [0336]). Aus der Perspektive des Fachmanns ist eine Content-Engine ein computergesteuertes System, das Inhalte bereitstellt und/oder verwaltet.

Unter einer Vorverarbeitung der analogen Mikrofonsignale (vgl. Merkmal **1.2.13** - „analog signals picked up by the microphone are pre-processed“) versteht das Streitpatent insbesondere eine Analog-Digital-Wandlung, Verschlüsselung, Kompression oder Konversion in eine andere Sprachdatenrepräsentation (Streitpatentschrift, Absatz [0064]).

5.1 Die Merkmale **1.2.12** bis **1.2.14** sind entweder aus Druckschrift **NK9** bekannt oder aus dieser unmittelbar ableitbar. Mit ihnen kann eine Patentfähigkeit des

Patentanspruchs 1 gemäß Hilfsantrag 5 vor dem Hintergrund der kombinierten Lehren der Druckschriften **NK11** und **NK9** nicht begründet werden.

Gemäß **NK9** stellen sowohl das „information distribution center 12“ als auch das Headend 125 dynamische, sich zeitlich ändernde Programmführerdaten - z.B. eine Liste verfügbarer Kanäle oder Programm- oder Künstlernamen - bereit, die von dem Spracherkennungssystem zur sprachgesteuerten Auswahl von Kanälen, Programmen oder Videos verwendet werden können (**NK9**, Spalte 5, Zeile 39 bis 67 - „terminal unit 16 receives vocabulary data from information distribution center 12“; Spalte 28, Zeile 50 bis Spalte 29, Zeile 22 sowie Table III - „Additional vocabulary data may be downloaded from head-end installation 125 [...] the vocabulary may be periodically updated to take into account changes in the offerings of a particular system [...] a user may navigate the electronic programming guide of Fig. 11 by saying „GOTO ESPN“; s. ferner Spalte 22, Zeile 54 bis 59 i. V. m. Spalte 29, Zeile 23 bis Spalte 30, Zeile 54 und Spalte 32, Zeile 12 bis 36 - „title addressable program selection“, „EPG information provider 114-3 provides the program title and description data [...] ISP 42 downloads the EPG text and vocabulary definitions to the subscriber terminal units [...] search of the EPG data to find occurrences of the programs having the spoken title [...] For example, if the user is tuned to television video, commanding „FIND STAR TREK“ will cause a search of the EPG data to be carried out [...] The user may then enter a command limiting the search, such as „ONLY MARILYN MONROE“ [...]).

In diesem Zusammenhang können diejenigen Komponenten des „information distribution center 12“ oder des Headends 125, die die Programmführerdaten verwalten und bereitstellen, als Content-Engine im Sinne von Merkmal **1.2.12** angesehen werden.

Auch wenn das Spracherkennungssystem an der Sendestation 10 angeordnet ist (s.o., Abschnitt **IV.4.1.2.1**), wird der Fachmann selbstverständlich dafür sorgen, dass es die dynamischen Programmführerdaten von dieser Content-Engine erhält. Denn andernfalls ließe sich keine Sprachsteuerung realisieren, die auf den dynamischen Daten basiert.

Somit ist Merkmal **1.2.12** aus der Lehre der **NK9** ableitbar.

Gemäß **NK9** werden die analogen Sprachsignale in der Fernbedienung in digitale Sprachsignale konvertiert und ggf. auch in eine spektrale Repräsentation transformiert (Spalte 15, Zeile 64 bis Spalte 16, Zeile 2 und 32 bis 38 i. V. m. Figuren 5 und 6, Blöcke 324 und 340). Die analogen Sprachsignale werden somit in der Fernbedienung vorverarbeitet.

Auch Merkmal **1.2.13** ist somit aus **NK9** bekannt.

Ferner lehrt **NK9**, dass die Fernbedienung die Äußerungen eines Teilnehmers an die Terminaleinheit in Form von Infrarot- oder Radiofrequenzsignalen übermittelt (Spalte 6, Zeile 33 bis 53; Spalte 12, Zeile 23 bis 26; Spalte 15, Zeile 64 bis Spalte 16, Zeile 5).

Damit nimmt die **NK9** auch Merkmal **1.2.14** vorweg.

5.2 Die Beklagte ist der Auffassung, die „information providers“ 114-n befinden sich nicht im gleichen Netzwerk wie das Headend 125 und lieferten zudem keine Inhaltsstatusinformationen, die Metainformationen über die bereits im TV/Video-Netzwerk befindlichen Daten seien. Merkmal **1.2.12** sei daher gegenüber der **NK9** neu.

Dem kann jedoch nicht beigetreten werden.

Denn zum einen kann das in den Figuren 2A bis 2C gezeigte System als ein einziges übergeordnetes und insbesondere auch drahtgebundenes Netzwerk angesehen werden, das die „information providers“ 114-n mit dem Headend 125 und den Terminaleinheiten 160 verbindet (vgl. **NK9**, Spalte 9, Zeile 18 bis 23; Spalte 11, Zeile 3 bis 6). Zum anderen können die Informationen, die von den Komponenten 114-n bereitgestellt werden, an das Headend 125 weitergegeben und von diesem an die Terminaleinheiten weitergeleitet werden (vgl. **NK9**, Spalte 8,

Zeile 44 bis Spalte 11, Zeile 6), so dass auch das Headend 125 als Content-Engine angesehen werden kann, zumal sich die „information providers“ 114-n auch an demselben Ort wie das Headend 125 befinden können (Spalte 9, Zeile 23 bis 26).

Ferner impliziert der Begriff Inhaltsstatusinformationen („content status information“) nicht, dass diese Informationen Metainformationen über Daten sein müssen, die sich bereits im Netzwerk befinden (s.o., Abschnitt **5**, Ausführungen zu Merkmal **1.2.12**). Zudem kann davon ausgegangen werden, dass Videofilme, die anhand eines Programm- oder Künstlernamens beispielsweise über einen Pay-Per-View-Dienst (vgl. **NK9**, Spalte 32, Zeile 12 bis 36 - „by speaking [...] “ONLY SYLVESTER STALLONE““) ausgewählt werden, in dem Kabelfernsehnetzwerk abspielbereit gespeichert sind.

5.3 Mit Blick auf die Ausführungen zum Patentanspruch 1 nach Hilfsantrag 4 beruht der Gegenstand des Patentanspruchs 1 gemäß Hilfsantrag 5 somit nicht auf erfinderischer Tätigkeit, so dass das Streitpatent in seiner Fassung nach Hilfsantrag 5 ebenfalls keinen Bestand hat.

6. Entsprechendes gilt für das Streitpatent in seiner Fassung nach **Hilfsantrag 6**, weil die gegenüber dem erteilten Patentanspruch 1 hinzugekommenen Merkmale ausgehend von Druckschrift **NK12** nahegelegt sind.

Patentanspruch 1 nach Hilfsantrag 6 unterscheidet sich vom erteilten Patentanspruch 1 darin, dass auf Merkmal **1.2.5** die Merkmale

- 1.2.15** wherein the speech recognition system (3200, 1330) communicates with a content engine (1340), indicating a current user location in a menu structure and/or a requested user action, in particular a video request;
- 1.2.16** and wherein the content engine (1340) provides at least one multimedia stream (1342) to a transceiver (1320) from which it is send to an individual user site.

folgen.

Gemäß Merkmal **1.2.15** soll das Spracherkennungssystem mit einer Content-Engine kommunizieren, indem es einen Ort in einer Menüstruktur und/oder eine „requested user action“ anzeigt. Unter einer solchen Aktion versteht der Fachmann eine Handlung des Teilnehmers, zu der dieser aufgefordert worden ist oder auch eine Anfrage des Teilnehmers, die gemäß dem Zusatz „in particular a video request“ insbesondere auf den Abruf eines Videos hindeuten kann („indicating [...] a requested user action“).

Merkmal **1.2.16** bringt zum Ausdruck, dass die Content-Engine mindestens einen Multimediestrom an einen Transceiver übermittelt, der diesen Datenstrom an einen einzelnen oder einen „individuellen“ Teilnehmerstandort („individual user site“) sendet. Dies muss nicht zwangsläufig bedeuten, dass der Multimediestrom ausschließlich an diesen (d.h. an einen „einzigsten“) Teilnehmerstandort gesendet wird, sondern kann auch nur implizieren, dass der Multimediestrom an einen Teilnehmerstandort gesendet wird, der sich nicht in der Nähe anderer Teilnehmerstandorte befindet (d.h. „einzeln“ liegt) oder der Anfrage des Teilnehmers entspricht (und deshalb „teilnehmerindividuell“ ist).

Ein Transceiver ist aus fachmännischer Sicht eine Vorrichtung, die mindestens eine Sendeeinrichtung („Transmitter“) und mindestens eine Empfangseinrichtung („Receiver“) umfasst.

6.1 Auch mit den Merkmalen **1.2.15** und **1.2.16** kann eine Patentfähigkeit nicht begründet werden.

Wie bereits in Abschnitt **IV.1.2.2** ausgeführt, ist der **NK12** zu entnehmen, dass das Spracherkennungssystem 28 der sprachgesteuerten Auswahl und Änderung von Fernsehbildern, Audiodaten oder sonstigen Daten dient, die an einen Teilnehmer über ein Kabelfernsehsystem übermittelt werden. Dazu werden als Reaktion auf die von dem Spracherkennungssystem interpretierten Sprachbefehle der

Fernsehsignalgenerator 29 und ein Teil des Kabelfernsehsystems 30 gesteuert. Die Sprachbefehle können daher als eine Steuerungsanfrage des Teilnehmers angesehen werden (vor dem Hintergrund der **NK12** beispielsweise eine Anfrage zur Auswahl bestimmter Fernsehbilder). Die Einheiten 29 und 30 sind netzwerkseitig angeordnet und in **NK12** die einzigen Einheiten, die die gewünschten Inhalte bereitstellen und an die Teilnehmer übermitteln können. Daher erkennt der Fachmann in den Einheiten 29 und 30 ein Headend des Kabelfernsehsystems - und somit auch eine Content-Engine.

Es ist selbstverständlich, dass das Spracherkennungssystem dabei den Inhalt der gesprochenen Anfrage des Teilnehmers an diese Content-Engine in Form einer Nachricht weiterleitet, die auf die auszuwählenden oder zu verändernden Inhalte Bezug nimmt und somit auf die Anfrage des Teilnehmers hindeutet.

Somit ist Merkmal **1.2.15** aus **NK12** ableitbar.

Falls der Teilnehmer bestimmte Fernsehbilder zur Übermittlung ausgewählt hat, werden diese von den Einheiten 29 und 30 - einem Headend (s.o.) - als Bilddatenstrom an die Kabelbox 25 des Teilnehmers übertragen. Dass ein Headend hierfür einen Transceiver verwendet, geht nicht über eine fachübliche Maßnahme hinaus (vgl. **NK16**, Seite 15, Figur 7; s. auch **NK9**, Spalte 10, Zeile 65 bis Spalte 11, Zeile 6 i. V. m. Figur 2B - „distribution network may include [...] one or more optical transmitters 140, one or more optical receivers 142“), welche eine erfinderische Tätigkeit nicht stützen kann.

Somit gelangt der Fachmann auch zu Merkmal **1.2.16**, ohne erfinderisch tätig zu werden.

6.2 Die Beklagte ist der Auffassung, die **NK12** zeige die Verarbeitungskette „Spracherkennungssystem → Content-Engine → Transceiver → Teilnehmer“ nicht. So sei die Verbindung zwischen dem Spracherkennungssystem 28 und dem Fernsehsignalgenerator 29 nicht beschrieben; unklar sei zudem, warum dieser nicht mit dem Kabelfernsehsystem 30 verbunden sei. Zudem kommuniziere der

Signalsplitter 27 nur upstream, so dass im Dunkeln bleibe, wie der Kanal zur Kabelbox 25 realisiert sei.

Diesen Argumenten kann ebenfalls nicht beigetreten werden.

So trifft es zwar zu, dass in **NK12** ein Datenaustausch zwischen den Einheiten 25 und 27 bis 30 nur ansatzweise beschrieben ist. Gleichwohl leitet der Fachmann aus **NK12** ab, dass die Komponenten 29 und 30 in Übereinstimmung mit den von dem Spracherkennungssystem 28 erkannten Sprachinhalten gesteuert werden, die als eine Anfrage zur Bereitstellung bestimmter Inhalte angesehen werden können (s.o., Abschnitt **IV.1.2.2**). In Verbindung mit dem Fachwissen des Fachmanns hinsichtlich des Datentransfers zwischen Set-Top-Boxen und Headends ergibt sich auch die oben genannte Verarbeitungskette in naheliegender Weise.

6.3 Angesichts der Ausführungen zum Hilfsantrag 1 erweist sich auch der Gegenstand des Patentanspruchs 1 gemäß Hilfsantrag 6 als nicht auf erfinderischer Tätigkeit beruhend. Mit seinem Patentanspruch 1 fällt der gesamte Hilfsantrag 6.

7. Hilfsantrag 7 kann nicht günstiger beurteilt werden, da der Gegenstand seines Patentanspruchs 1 ausgehend von Druckschrift **NK11** nahegelegt ist.

Der Patentanspruch 1 nach Hilfsantrag 7 geht aus dem erteilten Patentanspruch 1 hervor, indem das Merkmal

1.2.17 further comprising at least one of the steps of:
determining said associated user site (1100) from said recognized speech content;
determining said associated user site (1100) from said recognized speech content and a speaker identification library;
determining said associated user site (1100) from said recognized speech content and a speech recognition library; and

determining said associated user site (1100) from an identification within said speech channel.

nach Merkmal **1.2.5** angefügt wird.

Merkmal **1.2.17** fügt der Lehre von Patentanspruch 1 mindestens einen von vier Teilschritten hinzu, wobei der erste bzw. vierte Teilschritt mit dem zweiten bzw. fünften Teilschritt des Patentanspruchs 2 des Streitpatents übereinstimmt.

Wie in Abschnitt **III.4.2** ausgeführt (s.o.), können die Merkmale dieser beiden Teilschritte ausgehend von **NK11** keine erfinderische Tätigkeit begründen, so dass auch Merkmal **1.2.17** keine Patentfähigkeit stützen kann.

8. Auch **Hilfsantrag 8** kann nicht günstiger als Hilfsantrag 7 beurteilt werden, da sich der Gegenstand seines Patentanspruchs 1 ausgehend von Druckschrift **NK11** ebenfalls als naheliegend erweist.

Der Patentanspruch 1 nach Hilfsantrag 8 geht aus dem erteilten Patentanspruch 1 hervor, mit der Maßgabe, dass die letzte Alternative des Merkmals **1.2.17** gestrichen wird, d.h. Merkmal

1.2.18 further comprising at least one of the steps of:
determining said associated user site (1100) from said recognized speech content;
determining said associated user site (1100) from said recognized speech content and a speaker identification library;
determining said associated user site (1100) from said recognized speech content and a speech recognition library.

auf Merkmal **1.2.5** folgt.

Merkmal **1.2.18** fügt der Lehre von Patentanspruch 1 mindestens einen von drei Teilschritten hinzu, wobei der erste Teilschritt mit dem zweiten Teilschritt des Patentanspruchs 2 des Streitpatents übereinstimmt.

Wie in Abschnitt **III.4.2.2** ausgeführt (s.o.), können die Merkmale dieses Teilschritts ausgehend von **NK11** keine erfinderische Tätigkeit begründen, so dass auch die Aufnahme von Merkmal **1.2.18** in den Patentanspruch 1 zu keinem patentfähigen Gegenstand führt.

9. Aus diesen Gründen war das Streitpatent, das somit in keiner der durch die Beklagte verteidigten Fassungen Bestand hatte, im beantragten Umfang für nichtig zu erklären.

V.

Die Kostenentscheidung beruht auf § 84 Abs. 2 Satz 1 und Satz 2 Halbsatz 1 PatG i. V. m. § 91 Abs. 1 ZPO.

Die Entscheidung über die vorläufige Vollstreckbarkeit beruht auf § 99 Abs. 1 PatG i. V. m. § 709 Satz 1 und 2 ZPO.

VI.

Rechtsmittelbelehrung

Gegen dieses Urteil ist das Rechtsmittel der Berufung gemäß § 110 PatG statthaft.

Die Berufung ist innerhalb eines Monats nach Zustellung des in vollständiger Form abgefassten Urteils - spätestens nach Ablauf von fünf Monaten nach Verkündung - durch einen in der Bundesrepublik Deutschland zugelassenen Rechtsanwalt oder

Patentanwalt schriftlich beim Bundesgerichtshof, Herrenstraße 45a, 76133 Karlsruhe, einzulegen.

Die Berufungsschrift muss

- die Bezeichnung des Urteils, gegen das die Berufung gerichtet ist, sowie
- die Erklärung, dass gegen dieses Urteil Berufung eingelegt werde,

enthalten. Mit der Berufungsschrift soll eine Ausfertigung oder beglaubigte Abschrift des angefochtenen Urteils vorgelegt werden.

Auf die Möglichkeit, die Berufung nach § 125a PatG in Verbindung mit § 2 der Verordnung über den elektronischen Rechtsverkehr beim Bundesgerichtshof und Bundespatentgericht (BGH/BPatGERVV) auf elektronischem Weg beim Bundesgerichtshof einzulegen, wird hingewiesen (www.bundesgerichtshof.de/erv.html).

Hartlieb

Dr. Forkel

Hoffmann

Dr. Himmelmann

Dr. Städele