

G10L

SPEECH ANALYSIS TECHNIQUES OR SPEECH SYNTHESIS; SPEECH RECOGNITION; SPEECH OR VOICE PROCESSING TECHNIQUES; SPEECH OR AUDIO CODING OR DECODING

Definition statement

This place covers:

- Processing of speech or voice signals in general ([G10L 25/00](#)).
- Production of synthetic speech signals, text to speech systems ([G10L 13/00](#)).
- Recognition of speech ([G10L 15/00](#)).
- Lyrics recognition from a singing voice ([G10L 15/00](#)).
- Speaker identification, authentication or verification ([G10L 17/00](#)).
- Singer recognition from a singing voice (i.e. speaker recognition on a singing voice) ([G10L 17/00](#)).
- Analysis of speech signals for bandwidth compression or extension, bit-rate or redundancy reduction ([G10L 19/00](#)).
- Coding/decoding of audio signals for compression and expansion using analysis-synthesis, source filter models or psycho-acoustic analysis ([G10L 19/00](#)).
- Modification of speech signals, speech enhancement, source separation ([G10L 21/00](#)).
- Processing of the speech or voice signal to produce another audible or non-audible signal, e.g., visual or tactile, in order to modify its quality or its intelligibility ([G10L 21/00](#)).
- Noise reduction or echo cancellation in an audio signal ([G10L 21/00](#)).
- Speech or voice analysis techniques specially adapted to analyse or modify audio signals, where the audio signals do not necessarily include speech or voice, are also covered in subgroups ([G10L 21/00](#), [G10L 25/00](#)).

Relationships with other classification places

Classification should be generally directed to appropriate subclasses, e.g. [G06F](#), [H03M](#), for mathematical models for audio analysis in general.

Classification should be generally directed to appropriate subclasses, e.g. [G10K](#), [G10H](#), [H04R](#), [H04S](#) when audio productions or general audio analysis or processing are of relevance.

Telegraphic communication is covered in subclass [H04L](#).

Telephonic communication is covered in subclass [H04M](#).

References

Limiting references

This place does not cover:

Devices for the storage of speech signals	G11B
Static stores	G11C
Encoding of compressed speech signals for transmission or storage	H03M 7/30

Application-oriented references

Examples of places where the subject matter of this place is covered when specially adapted, used for a particular purpose, or incorporated in a larger system:

Information retrieval of audio data	G06F 16/60
-------------------------------------	----------------------------

Application-oriented references

Broadcasting arrangements of audio	H04H 60/58
Name dialling controlled by voice recognition	H04M 1/271
Automatic arrangements for answering calls	H04M 1/64

References out of a residual place

Examples of places in relation to which this place is residual:

Acoustics not otherwise provided for	G10K 15/00
--------------------------------------	----------------------------

Informative references

Attention is drawn to the following places, which may be of interest for search:

Larynxes or tracheae prosthesis implantable into the body	A61F 2/20
Input/output arrangements for on-board computers	G01C 21/36
Measurement of sound waves in general	G01H
Direction-finders for determining the direction from which infrasonic, sonic or ultrasonic waves, not having a directional significance, are being received	G01S 3/80
Systems using the reflection or reradiation of acoustic waves	G01S 15/00
Sound input/output for computers	G06F 3/16
Compilation or interpretation of high level programme languages	G06F 8/41 , G06F 9/455
Information retrieval; Database structures therefor	G06F 16/00
Complex mathematical functions	G06F 17/10
General pattern recognition	G06F 18/00
Digital data processing methods or equipment specially adapted for handling, processing or translating natural language data	G06F 40/00
Image data processing	G06T
Arrangements for image or video recognition or understanding	G06V 10/00
Character recognition; Recognising digital ink; Document-oriented image-based pattern recognition	G06V 30/00
Recognition of biometric, human-related or animal-related patterns in image or video data	G06V 40/00
Individual entry or exit registers	G07C 9/00
Arrangements for influencing the relationship between signals at input and output, e.g. differentiating, delaying	G08C 13/00
Teaching speaking	G09B 19/04
Teaching or communicating with the blind, deaf or mute	G09B 21/00
Electronic musical instruments	G10H
Sound producing devices other than musical instruments or loudspeakers	G10K
Methods or devices for protecting against, or for damping, noise or other acoustic waves	G10K 11/16
Signal processing for recording	G11B 20/00
Error detection or correction in digital recording or reproducing; Testing involved in digital recording or reproducing	G11B 20/18

Editing; Indexing; Addressing; Timing or synchronising; Monitoring	G11B 27/00
Electronic circuits for sound generation	H03B
Amplifiers	H03F
Amplifiers using amplifying element consisting of two mechanically- or acoustically-coupled transducers, e.g. telephone-microphone amplifier	H03F 13/00
Gain control in amplifiers or frequency changers	H03G 3/00
Electronic filters	H03H
Coding, decoding or code conversion, error protection in general	H03M
Transmission	H04B
Means associated with receiver for limiting or suppressing noise or interference	H04B 1/10
Details of transmission systems, not characterised by the medium used for transmission, for reducing bandwidth of signals	H04B 1/66
Transmission systems employing ultrasonic, sonic or infrasonic waves	H04B 11/00
Transmission systems not characterised by the medium used for transmission characterised by the use of pulse modulation	H04B 14/02
Broadcast distribution systems	H04H
Time-division multiplex systems in which the transmission channel allotted to a first user may be taken away and re-allotted to a second user if the first user becomes inactive	H04J 3/17
Secret communication	H04K 1/00
Transmission of digital information, e.g. telegraphic communication	H04L
Telephonic communication	H04M
Arrangements of transmitters, receivers or complete sets to prevent eavesdropping, to attenuate local noise or to prevent undesired transmission; Special mouthpieces or receivers therefor	H04M 1/19
Arrangements for preventing acoustic feedback in telephonic communication	H04M 1/20
Devices for calling a subscriber whereby a plurality of signals may be stored simultaneously	H04M 1/27
Substation equipment, e.g. for use by subscribers including speech amplifiers	H04M 1/60
Automatic arrangements for answering calls	H04M 1/64
Interactive information services, e.g. directory enquiries	H04M 3/493
Simultaneous speech and telegraphic or other data transmission over the same conductors	H04M 11/06
Systems for transmission of a pulse code modulated video signal with one or more other pulse code modulated signals, e.g. an audio signal, a synchronising signal	H04N 7/52
Switching systems	H04Q
Loudspeakers, microphones, gramophone pick-up or like acoustic electromechanical transducers; Deaf-aid sets; Public address systems	H04R
Stereophonic arrangements	H04R 5/00
Public address systems	H04R 27/00

Stereophonic systems, e.g. spatial sound capture, matrixing of audio signals in the decoded state	H04S
---	----------------------

Glossary of terms

In this place, the following terms or expressions are used with the meaning indicated:

speech	definite vocal sounds that form words to express thoughts and ideas.
voice	sounds generated by vocal chords or synthetic versions thereof.
audio signal	of or relating to humanly audible sound, meant to include speech, voice, music, silence or background noise, or any combinations thereof.

Synonyms and Keywords

In patent documents, the following abbreviations are often used:

AAC	Advanced Audio Coding
ACELP	Algebraic Code Excited Linear Prediction
ADPCM	Adaptive Differential Pulse Code Modulation
AMR, AMR-NB	Adaptive Multi-Rate
AMR-WB	Adaptive Multi-Rate Wideband
ANN	Artificial Neural Network
AR	Autoregressive
ASR	Automatic Speech Recognition
BLP	Backward Linear Prediction
BP	Back Propagation
BSAC	Bit Sliced Arithmetic Coding (audio coding from MPEG-4 Part 3)
CELP	Code Excited Linear Prediction
DCT	Discrete Cosine Transform
DFT	Discrete Fourier Transform
DPCM	Differential Pulse Code Modulation
DRM	Digital Rights Management
DTX	Discontinuous Transmission
EVRC, EVRC-B	Enhanced Variable Rate CODEC
FFT	Fast Fourier Transform
FIR	Finite Duration Impulse Response
FLP	Forward Linear Prediction
GMM	Gaussian mixture model
HMM	Hidden Markov model
HVXC	Harmonic Vector eXcitation Coding
IDCT	Inverse Discrete Cosine Transform
IVR	Interactive Voice Response

LMS	Least Mean Square
LPC	Linear Predictive Coding
LSF	Line Spectral Frequencies
LSP	Line Spectral Pairs
LTP	Long Term Prediction
MBE	Multi-Band Excitation
MDCT	Modified Discrete Cosine Transform
MELP	Mixed Excitation Linear Prediction
MLP	Multi-Layer Perceptron
MP3	MPEG1 or MPEG2 audio layer III
MPEG	Motion Picture Experts Group
MPEG 1 audio	Standard ISO/IEC 11172-3
MPEG 2 audio	Standard ISO/IEC 13818-3
MPEG 4 audio	Standard ISO/IEC 14496-3
MPEG 21	Standard ISO/IEC 21000
MSE	Mean Square Error
NB – WB	Narrowband – Wideband
PARCOR	Partial Correlation
PWI	Prototype Waveform Interpolation
REL P	Residual Excited Linear Prediction
SBR	Spectral Band Replication
TDNN	Time Delay Neural Network
TTS	Text-to-Speech
USAC	Unified Speech and Audio Coding
VoIP	Voice over Internet Protocol
VLSR	Very Large Speech Recognition
VQ	Vector Quantization
VSELP	Vector Sum Excited Linear Prediction
V/UV	Voiced/Unvoiced
VXML or VoiceXML	W3C's standard XML format

G10L 13/00

Speech synthesis; Text to speech systems

Definition statement

This place covers:

- synthesis of speech from text, concatenation of smaller speech units, grapheme to phoneme conversion;
- modification of the voice for speech synthesis: gender, age, pitch, prosody, stress.
- hardware or software implementation details of a speech synthesis system

References

Application-oriented references

Examples of places where the subject matter of this place is covered when specially adapted, used for a particular purpose, or incorporated in a larger system:

Speech synthesis in games	A63F 9/24
Navigation systems for vehicles, guidance using speech synthesis.	G01C 21/3629
Electronic musical instruments	G10H
Sound producing other than musical instruments	G10K
Electric switches with speech feedback	H01H
Speech synthesis in mobile phones	H04M 1/00

Informative references

Attention is drawn to the following places, which may be of interest for search:

Excitation coding of a speech signal	G10L 19/08
Processing or translation of natural language	G06F 40/40

Synonyms and Keywords

In patent documents, the following abbreviations are often used:

HMM	Hidden Markov Model
TTS	Text To Speech

G10L 13/027

Concept to speech synthesisers; Generation of natural phrases from machine-based concepts (generation of parameters for speech synthesis out of text [G10L 13/08](#))

Definition statement

This place covers:

concepts used for speech synthesis can be linked to an emotion to be conveyed (US2010329505), a communication goal driving a dialogue (US2010241420), image-to-speech (US2010231752), native sounding speech (US2004030554)

References

Limiting references

This place does not cover:

Processing or translation of natural language	G06F 40/40
---	----------------------------

G10L 15/00**Speech recognition ([G10L 17/00](#) takes precedence)****Definition statement**

This place covers:

- recognition of text or phonemes from a spoken audio signal;
- spoken dialog interfaces, human-machine spoken interfaces
- topic detection in a dialogue, semantic analysis, keyword detection, spoken command and control
- context dependent speech recognition (location, environment, age, gender, etc.)
- parameter extraction, acoustic models, word models, grammars, language models for speech recognition
- recognition of speech in a noisy environment
- recognition of speech using visual clues
- feedback of the recognition results, disambiguation of speech recognition results
- dedicated hardware or software implementations, parallel and distributed processing of speech recognition engines

References**Application-oriented references**

Examples of places where the subject matter of this place is covered when specially adapted, used for a particular purpose, or incorporated in a larger system:

Spoken command and control of surgical instruments	A61B 17/00
Speech input in video games	A63F 13/00
Voice control for systems within a vehicle	B60R 16/0373
Speech input for vehicle navigation systems	G01C 21/3608
Sound input arrangements for computers	G06F 3/16
Teaching how to speak	G09B 19/04
Name dialling controlled by voice recognition	H04M 1/271
Speech interaction details in automatic or semi-automatic exchange systems for interactive information services	H04M 3/4936

Informative references

Attention is drawn to the following places, which may be of interest for search:

Information retrieval of audio data	G06F 16/60
Complex mathematical functions	G06F 17/10
Pattern recognition	G06F 18/00
Handling natural language data	G06F 40/00
Image or video recognition or understanding	G06V
Face recognition, lip reading without acoustical input	G06V 40/16
Educational appliances	G09B 5/06
Signal processing for recording	G11B 20/00

Synonyms and Keywords

In patent documents, the following abbreviations are often used:

ANN	Artificial neural network
ASR	Automatic speech recognition
CSR	Continuous speech recognition
GMM	Gaussian mixture model
HMM	Hidden Markov model
IVR	Interactive voice response
MLP	Multi layer perceptron
VLSR	Very large speech recognition

G10L 17/00

Speaker identification or verification techniques

Definition statement

This place covers:

- Recognition, identification of a speaker
- Verification, authentication of a speaker
- Preprocessing operations, e.g. segment selection; Pattern representation or modelling, e.g. based on linear discriminant analysis [LDA], principal components; Feature selection or extraction
- Dialog, prompts, passwords for identification
- Training, model building or enrolment
- Decision making techniques, pattern matching strategies
- Multimodal identification including voice
- Hidden Markov Models
- Artificial neural networks, connectionist approaches
- Pattern transformations and operations aimed at increasing system robustness, e.g. against channel noise, different working conditions
- Identification in noisy condition
- Interactive procedures, man-machine interface, e.g. user prompted to utter a password or predefined text
- Recognition of special voice characteristics, e.g. for use in a lie detector; recognition of animal voices
- Imposter detection

References

Informative references

Attention is drawn to the following places, which may be of interest for search:

Information retrieval of audio data	G06F 16/60
Complex mathematical functions	G06F 17/10
Pattern recognition	G06F 18/00
User authentication in security arrangements for restricting access by using biometric data, e.g. voice prints	G06F 21/32
Machine learning	G06N 20/00

Arrangements for image or video recognition or understanding	G06V 10/00
Recognition of biometric, human-related or animal-related patterns in image or video data	G06V 40/00
Individual entry or exit registers, access control with identity check using personal physical data	G07C 9/25
Secret secure communication including means for verifying the identity or authority of a user	H04L 9/32
Secret secure communication including means for verifying the identity or authority of a user	H04L 9/32
Centralised arrangements for answering calls; Centralised arrangements for recording messages for absent or busy subscribers	H04M 3/50

Glossary of terms

In this place, the following terms or expressions are used with the meaning indicated:

speaker verification or authentication	refers to verifying that the user claimed identity is real, he is otherwise an impostor. Speaker recognition, or identification, aims at determining who the user is among a closed (finite number) set of users. He is otherwise unknown.
impostor	someone actively trying to counterfeit someone else's identity.

Synonyms and Keywords

In patent documents, the following abbreviations are often used:

ANN	Artificial neural network
ASR	Automatic speech recognition
GMM	Gaussian mixture model
HMM	Hidden Markov model
IVR	Interactive voice response
MLP	Multilayer perceptron

In patent documents, the following words/expressions are often used as synonyms:

- "goat" and "sheep"
- "predator" and "wolf"

In patent documents, the following words/expressions are often used with the meaning indicated:

A goat or sheep	often refers to a person whose voice is easy to counterfeit.
A wolf or predator	often refers to a person who can easily counterfeit someone else's voice or is often identified as someone else.

G10L 19/00

Speech or audio signals analysis-synthesis techniques for redundancy reduction, e.g. in vocoders; Coding or decoding of speech or audio signals, using source filter models or psychoacoustic analysis (in musical instruments [G10H](#))

Definition statement

This place covers:

Techniques for the reduction of data from audio sources, i.e. compression of audio. These techniques are applied to reduce the quantity of information to be stored or transmitted, but are independent of the end-application, medium or transmission channel, i.e. do only exploit the properties of the source signal itself or the final receiver exposed to this signal (the listener).

Mainly two types of sources can be distinguished :

"speech only" encompass signals produced by human speakers, and historically was to be understood as mono-channel, single speaker "telephone quality" speech having a narrow bandwidth limited to max. 4kHz. Encoding of speech only sources primarily aim at reducing the bit-rate while still providing fair intelligibility of the spoken content, but not always fidelity to the original.

"Audio signal" is broader and comprises speech as well as background information, e.g. music source having multiple channels. Encoding of audio deals primarily with transparent, i.e. "high fidelity" reproduction of the original signal.

The compression techniques can also be distinguished as being :

Lossy or Lossless, i.e. whether a perfect reconstruction of the source is possible, or only a perceptually acceptable approximation can be done.

The techniques classified in this subclass are based either on modelling the production of the signal (voice) or the perception of it (general audio).

References

Limiting references

This place does not cover:

Coding of signals within electronic musical instruments	G10H
---	----------------------

Informative references

Attention is drawn to the following places, which may be of interest for search:

Complex mathematical functions	G06F 17/10
Signal processing not specific to the method of recording or reproducing	G11B 20/00
Editing; Indexing; Addressing; Timing or synchronizing; Monitoring;	G11B 27/00
Compression	H03M 7/30
Detecting, preventing errors in received information	H04L 1/00
Quality monitoring in automatic, semi automatic exchanges	H04M 3/2236
Quality control of voice transmission between switching centres	H04M 3/2236
Simultaneous speech and data transmission	H04M 11/06
Transmission of audio and video in television systems	H04N 7/52

Stereophonic arrangements	H04R 5/00
Stereophonic systems	H04S
Wireless communication networks	H04W

Glossary of terms

In this place, the following terms or expressions are used with the meaning indicated:

audio signal	is meant to include speech, music, silence or background signal, or any combinations thereof, unless explicitly specified
--------------	---

Synonyms and Keywords

In patent documents, the following abbreviations are often used:

CELP	Code Excited Linear Prediction
CTX	Continuous transmission
DTX	Discontinuous transmission
HVXC	Harmonic Vector eXcitation Coding
LPC	linear prediction coding
MBE	Multiband Excitation
MELP	Mixed Excitation Linear Prediction
MOS	mean opinion score
MPEG	Moving Picture Experts Group
MPEG1 audio	Standard ISO/IEC 11172-3
MPEG2 audio	Standard ISO/IEC 13818-3
MPEG4 audio	Standard ISO/IEC 14496-3
MP3	MPEG 1 Layer III
PCM	pulse code modulation
PWI	Prototype Waveform Interpolation
SBR	Spectral Band Replication

In patent documents, the following words/expressions are often used as synonyms:

- "perceptual" and "psychoacoustic"

G10L 19/24

Variable rate codecs, e.g. for generating different qualities using a scalable representation such as hierarchical encoding or layered encoding

Definition statement

This place covers:

Coding of a signal with rate adaptation, e.g. adapted to voiced speech, unvoiced speech, transitions and noise/silence portions.

Definition statement

Coding of a signal with a core encoder providing a minimum level of quality, and extension layers to improve the quality but requiring a higher bitrate. It includes parameter based bandwidth extension (i.e. SBR) or channel extension.

This group is in opposition to [G10L 21/038](#) in which the bandwidth extension is artificial, i.e. based on the only narrowband encoded signal.

References

Informative references

Attention is drawn to the following places, which may be of interest for search:

Artificial bandwidth extension, i.e. based on the only narrowband encoded signal	G10L 21/038
Spatial sound recording	H04R 5/00
Spatial sound reproduction	H04S

G10L 21/00

Speech or voice signal processing techniques to produce another audible or non-audible signal, e.g. visual or tactile, in order to modify its quality or its intelligibility ([G10L 19/00](#) takes precedence)

Definition statement

This place covers:

The subgroup deals with speech or voice modification applications, but receives also applications for speech or voice analysis techniques specially adapted to analyse or modify audio signals not necessarily including speech or voice but which are not music signals ([G10H](#)).

- Speech or voice modification applications, but receives also applications for speech or voice analysis techniques specially adapted to analyse or modify audio signals, where the audio signals do not necessarily include speech or voice but which are not music signals ([G10H](#)).
- Bandwidth extension of an audio signal.
- Improvement of the intelligibility of a coded speech signal.
- Removal of noise from an audio signal.
- Removal of echo from an audio signal.
- Separation of audio sources.
- Pitch, speed modification of an audio signal.
- Voice morphing.
- Visualisation of audio signals (e.g. sonograms).
- Lips or face movement synchronisation with speech (e.g. phonemes - visemes alignment).
- Face animation synchronisation with the emotion contained in the voice or speech signal.

References

Limiting references

This place does not cover:

Speech or audio signal analysis-synthesis techniques for redundancy reduction, e.g. in vocoders; Coding or decoding of speech or audio signals, using source-filter models or psychoacoustic analysis	G10L 19/00
---	----------------------------

References out of a residual place

Examples of places in relation to which this place is residual:

Electronic musical instruments	G10H
Loudspeakers, microphones, gramophone pick-up or like acoustic electromechanical transducers; deaf-aid sets; public address systems	H04R
Stereophonic systems	H04S

Informative references

Attention is drawn to the following places, which may be of interest for search:

Direction finder	G01S 3/00
Complex mathematical functions	G06F 17/10
3D Animation	G06T 13/20
Animation based on audio data, talking heads	G06T 13/205
Signal processing not specific to the method of recording or reproducing	G11B 20/00
Signal processing not specific to the method of recording or reproducing, for reducing noise	G11B 20/24
Editing; Indexing; Addressing; Timing or synchronizing; Monitoring	G11B 27/00
Gain control in amplifiers where the control is dependent upon ambient noise level or sound level	H03G 3/32
Reducing echo effect or singing in line transmissions systems	H04B 3/20
Transmission systems not characterised by the medium used for transmission using pulse code modulation, e.g. for reducing noise or bandwidth	H04B 14/04
Reducing noise or bandwidth in transmission systems not characterised by the medium used for transmission	H04B 14/046
Echo suppression in hand-free telephones	H04M 9/08
Hearing aids	H04R 25/00
Public address systems	H04R 27/00

Glossary of terms

In this place, the following terms or expressions are used with the meaning indicated:

viseme	a visual representation of the mouth, lips, tongue and teeth corresponding to a phoneme.
--------	--

Synonyms and Keywords

In patent documents, the following abbreviations are often used:

BSS	Blind source separation
LDA	Linear discriminant analysis
NB	Narrowband
PCA	Principal component analysis
SBR	Spectral Band Replication

WB	Wideband
----	----------

G10L 21/0356

for synchronising with other signals, e.g. video signals

Definition statement

This place covers:

Visemes are selected to match with the corresponding speech segment, or the speech segments are adapted/chosen, to match with the viseme. This symbol also encompasses the coarticulation effects as used in facial character animation or talking heads.

References

Informative references

Attention is drawn to the following places, which may be of interest for search:

Facial character animation per se	G06T 13/205
-----------------------------------	-----------------------------

G10L 21/038

using band spreading techniques

Definition statement

This place covers:

Bandwidth extension taking place at the receiving side, e.g. generation of artificial low or high frequency components, regeneration of spectral holes, based on the only narrowband encoded signal. This is in opposition with [G10L 19/24](#) wherein parameters are computed during the encoding step to enable bandwidth extension at the decoding step.

References

Informative references

Attention is drawn to the following places, which may be of interest for search:

Parameter based bandwidth extension (e.g. SBR)	G10L 19/24
--	----------------------------

G10L 25/00

Speech or voice analysis techniques not restricted to a single one of groups [G10L 15/00](#) - [G10L 21/00](#) (muting semiconductor-based amplifiers when some special characteristics of a signal are sensed by a speech detector, e.g. sensing when no signal is present, [H03G 3/34](#))

Definition statement

This place covers:

- Processing of speech or voice signals in general, in particular detection of a speech signal, end points detection in noise, extraction of pitch, measure of the voicing, emotional state, voice pathology or other speech or voice related parameters.
- Extracted parameters, e.g. techniques for evaluating correlation coefficients, zero crossing, prediction coefficients or formant information.

Definition statement

- Analysis technique, e.g. neural network, fuzzy, chaos, genetic algorithm or coding technique.
- Analysis window (window function).
- Specially adapted for particular use, e.g. for comparison and discrimination, evaluating synthetic and decoded voice signals, for transmitting result of analysis.
- Speech or voice analysis techniques specially adapted to analyse audio signals, where the analysed audio signals do not necessarily include speech or voice, such as audio scene segmentation, jingle detection, separation from music or noise or detection of particular sounds.
- Modeling vocal tract parameters.
- Detection of presence or absence of speech signals.
- Pitch determination of speech signals.
- Discriminating between voiced and unvoiced parts of speech signals.

References

Limiting references

This place does not cover:

Muting semiconductor-based amplifiers when some special characteristics of a signal are sensed by a speech detector, e.g. sensing when no signal is present	H03G 3/34
---	---------------------------

Informative references

Attention is drawn to the following places, which may be of interest for search:

Comfort noise	G10L 19/012
Karaoke or singing voice processing, parameter extraction for musical signal categorisation, electronic musical instruments	G10H
Gain or frequency control	H03G 3/342
DTX communication	H04J 3/175
Switching of direction of transmission by voice frequency in two-way loud-speaking telephone systems	H04M 9/10
Multiplex systems	H04Q 1/30

Glossary of terms

In this place, the following terms or expressions are used with the meaning indicated:

audio signal	of or relating to humanly audible sound, e.g. it comprises any combination of background noise or silence, voice or speech, music
--------------	---