(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

PCT

(43) International Publication Date
10 April 2008 (10.04.2008)

(10) International Publication Number
WO 2008/041878 A2

(51) International Patent Classification: Not classified

(21) International Application Number: PCT/RS2007/000017

(22) International Filing Date: 19 September 2007 (19.09.2007)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
P-2006/0551 4 October 2006 (04.10.2006) RS

(71) Applicant: **MICRONAS NIT** [RS/RS]; Fruskogorska 1 la, 21000 Novi Sad (RS).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **SARIC, Zoran** [RS/RS]; Vukasovica 65/7, 11000 Novi Beograd (RS). **JOVICIC, Slobodan** [RS/RS]; Visnjicki venae 67, 11000 Beograd (RS). **KOVACEVIC, Vladimir** [RS/RS]; Radnicka 35A, 21000 Novi Sad (RS). **TESLIC, Nikola** [RS/RS]; Bul. Cara Lazara 29, 21000 Novi Sad (RS). **KUKOLJ, Dragan** [RS/RS]; Narodnog fronta 31, 21000 Novi Sad (RS).

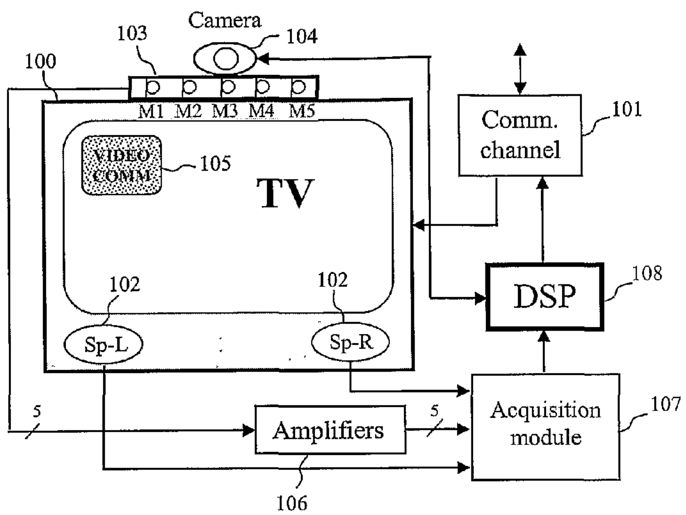
(81) **Designated States** *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW

(84) **Designated States** *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, **DK**, EE, ES, FI, FR, GB, GR, HU, IE, **IS**, **IT**, LT, LU, LV, MC, MT, NL, PL, PT, RO, SE, **SI**, SK, TR), OAPI (BF, **BJ**, CF, CG, CI, CM, GA, GN, GQ, GW, ML, **MR**, NE, SN, TD, TG).

Declaration under Rule 4.17:
— *of inventorship (Rule 4.17(iv))*

Published:
— *without international search report and to be republished upon receipt of that report*

(54) **Title:** SYSTEM AND PROCEDURE OF FREE SPEECH COMMUNICATION USING A MICROPHONE ARRAY



(57) **Abstract:** The invention relates to the system and procedure for hand-free voice communication in video-phone or teleconference using a microphone array, whose main purpose is to make a quality recording of speaker in room, in the situation of larger expansion, with presence noise, with acoustic echo, produced by distance speaker and TV program, room reverberation and movement of the speaker in room. System contains: digital TV receiver and digital camera for picture reproduction and shooting, respectively, stereo loudspeakers and microphone array for sound reproduction and recording, respectively, amplifier and acquisition module for audio signals and DSP for acoustic signal processing. The procedure for microphone signal processing is done in frequency domain and it contains: acoustic echo suppression made of two signals: far-end speaker signal and stereo TV signal, acoustic spatial filtering of near-end speaker in accordance with noise sources and room reverberation, based on adaptive characteristic of microphone array directivity, of speaker localization in horizontal plane, of suppression of all residual noises and adaptive gain control of transmitting signal.

WO 2008/041878 A2

SYSTEM AND PROCEDURE OF FREE SPEECH COMMUNICATION USING A MICROPHONE ARRAY

Technical Field

The invention belongs to the field of acoustic signal processing, precisely speaking to the methods of acoustic echo cancellation, location and selection of an active speaker in the presence of a reverberations in the acoustic environment and the noise suppression by means of microphone array.

Background Art

Hands-free full-duplex speech communication systems are used in many existing applications, such as: video-phone systems, teleconference systems, room and car hands-free systems, human-machine interface using voice, etc.

Usage of the hands-free speech communication systems implies not specified talker position in the acoustic environment, with variable distances from system's microphones and loudspeakers. The hands-free speech communication in such unknown conditions is reason for the number of technical problems, which should be solved, in order to preserve good quality of the speech communication.

Basic problem is acoustic echo generated by partial acoustic energy transmission from a loudspeaker to the microphone, so the speaker on far-end is able to hear his own voice as an obstruction. Conventionally, signal echo canceling is done by adaptive filter using estimation of transfer function of acoustic echo between loudspeaker and microphone, so that its exit gets approximately same signal as acoustic echo signal. Deduction two of these signals cancels acoustic echo. However, canceling echo can not be perfect because of systems non-linearity and acoustics ambience non-steady. As a result it shows residual echo signal. At that basic request stays, recorded speech signal of near-end shouldn't be exposed by echo suppression and its process.

In the acoustic ambient, acoustic disturbances of different nature and causes may appear. Those disturbances could be stationary and non-stationary (for example: computer noises or car noise) and they come from many different sources located on different positions in the room or space where the speaker stands.

Besides that, in closed rooms (as a work rooms, halls and automobile-cabins) it shows up the effect of reverberation as an after effect of multiple acoustic wave reflections from walls and obstacles. Since the acoustic ambient besides the speaker contain sources of disturbances, the desired signal (coming from the speaker) must be separated from the disturbances in order to make possible its own recording. Conventionally, this problem may be solved by using a microphone array having a number of microphones ordered in line at minimum inter-distance. With appropriate processing of microphone array signal, direction dependent sensitivity of microphone system may be achieved. Such microphone systems has narrow directivity characteristic, enough to record only the actual speaker in the acoustic ambient, while the signals of dislocated noise sources are suppressed, thereby providing higher signal-to-disturbance ratio. The gain depends on: directivity of the microphone array (width of the main lobe), side-lobe size, separability of speech sources and noise sources (to close sources are difficult to separate), reverberation time, non-stationary acoustic sources, etc.

Determination of speaker direction in acoustic ambient and steering the diiectivity of microphone array according toward it is an important problem in hands-free communication systems. The procedures of determining the speaker direction 'are very sensitive to disturbances present in the ambient, specially: to non-stationary speaker (if it moves within ambient) and if there are several speakers in a given ambient simultaneously speaking (cocktail party effect). The determination of relative direction of the actual speaker to the microphone array in horizontal plane (determination of azimuth), is very important step in video-phone and teleconferencing systems, because of need to determine the speaker coordinates which are used for moveable camera control in the system.

During speech recording in an acoustic ambient, the problem of additive stationary or non-stationary noise always appears so as the residual noise in processing of acoustic signals. They degrade the quality of the recorded speech signal. If they are intense enough, they may even reduce the perspicuity of the speech. There are many algorithms for noise reduction (NR), optimized for specific noise types. The common requirement for all of them is to improve the signal to noise ratio, but to avoid distorting of speech signal and reduction of its perspicuity.

Variable ambient conditions, and variable distance between the speaker and microphone array, require automatic gain control (AGC), which makes the speaker voice level constant and more comfort for the receiver at the far-end of the communication channel. Automatic gain control in full-duplex systems requires additional information from near-end speech activity detector, from far-end speech activity detector and acoustic echo canceller.

Refer to above mentioned technical problems in solution of "hand-free" communication system for speech signal transmission in full-duplex and its usage in video-phone and/or teleconference systems, are very complex. Those problems demand one integral and optimal solution approach, considering real time system operation based on commercial platform of digital signal processor (DSP).

Quality of speech recording in the presence acoustic noises and room reverberations made a complex problem. In the conditions when the useful speech signal spectrum are overlapping with presence noises spectrum, using a single channel processing it is not possible to improve significantly of speech signal quality. In accordance with digital signal processing development and purchasing of enough powerful computer power of DSP, a way of multi-microphone procedure applying acoustic signals processing is open. Benefits of microphone array in relation to single channel processing is adaptation capability of its spatial receipt characteristics (directivity characteristic) to instantly schedule of chosen speaker and define noises in room. At that point, they realize a maximum suppression of presence noises, at the same time the speaker is emphasized. Main problems by microphone arrays usage are (M.S.Brandstein, D.B. Ward (Eds.), *Microphone Arrays: Signal Processing Techniques and*

Applications, Springer, Berlin 2001; Y. Huang, J. Benesty, *Audio signal processing for next generation multimedia communication systems*, Kluwer Academic Publ.; 2004): chosen speaker exactly location outset, outset of exactly number and positions of room presence noises, multi-reflections of useful source and noise of the room walls and non-steady of acoustic noise sources and chosen speaker.

When the microphone array is used in video-phone or teleconference systems, in full duplex function, than the number of possible problems is getting larger. The biggest problem is presence of acoustic echo, and then need for automatic gain control (AGC) of system transmitter part, as well as possible presence of system non-steady, called microphony. Additional problem, which is being observed in this patent, is presence of TV program signal, which shows up as an additive acoustic echo on entrance of microphone array.

Large number of mentioned problems has been generated and made very different kind of solutions, which has been patented and which could solve some of problems or few integral problems. For example: U.S. published patent application 2006/ 0153360 Al, filled September 2nd 2005, entitled "Speech signal processing with combined noise reduction and echo compensation", gives integral solution of echo reduction and noise reduction, then U.S. published patent application 7,035,415 B2, filled May 15th 2001, entitled "Method and device for acoustic echo cancellation combined with adaptive beamforming", which gives integral solution of echo reduction and forming of directed microphone array characteristic, then EP published patent application 1 633 121 Al, filled September 3rd 2004, entitled "Speech signal processing with combined adaptive noise reduction and adaptive echo compensation", gives integral solution of residual echo reduction and noise reduction, then EP published patent application 1 571 875 A2, filled February 23rd 2005, entitled " A system and method for beamforming using a microphone array", which gives solution for only directed microphone array characteristic forming, then EP published patent application 1 581 026 Al, filled March 17th 2004, entitled "Method for detecting and reducing noise from a microphone array" gives solution only for noise reduction in microphone array, as well as EP published patent application 1 286 175 A2, filled August 1st 2002, entitled "Robust talker localization in reverberant environment", gives solution only for talker localization in reverberant room.

Integral solution all mentioned problems, realized in this patent, join positive characteristics of particular signal processing of mentioned problems and their solutions, they are going to be solved integrally in frequency domain, optimizing computer resources and gives real time solutions, securing quality of free speech communication in video-phone and/or teleconference systems.

Disclosure of the Invention

Subject of this patent is free speech communication system in video-phone or teleconference applying, which use microphone array and complex acoustic signal processing, which should secure better quality and clearness of speech signal in complex acoustic ambience, in which many previous mentioned failures are separately or integral eliminated.

System, which is subject of this patent, transmits speech and as transmitting medium is being used digital television. For recording and reproduction of speech signal is being used microphone array and loudspeaker, respective, which are integral TV receiver components. When we talk about video-phone or teleconference applying for recording and picture reproduction than we use digital camera and respective digital TV receiver.

Invention essence is specific processing of speech signal, which has been recorded in one acoustic ambience in room where the speaker and system are present. For recording of speaker in room, which stands on define distance (few meters distance) from TV receiver, system uses microphone array of N microphones. Microphone array records all present room signals: useful signal as a directed wave, which gets from the talker to the microphone and different noise signals. As noise signals it shows up: acoustic echo as one loudspeaker direct wave, which is emitting interlocutor voice from the far-end of communication channel, acoustic echo as a directly sound wave, which are emitting stereo TV program, direct waves taken from one or more source of noise or also other sources, which we can hear in the room and reflected waves (room echo), made by their own sources of noise, including speaker, and all those noise, which appear to show during the room reverberation. We should emphasis that noise sources in the room can be stationary or non-steady, which is frequently matter, as by its characteristics, so as by its room location (mobile sound sources).

Different kinds of noises required different techniques for its eliminating, and this invention essence is one optimally designed algorithm, which should at most eliminate all noises and which should secure the best speech signal quality, which is going to be transmitted to the interlocutor on the far-end of communication channel.

Microphone signals from microphone array are being processed in one digital form in DSP, completely in one frequency domain. This domain enables certain advantages, as a processing speed and computer operation number, which is very important for DSP and its real time work. For acoustic echo cancellation it is necessary to put in all loudspeaker signals into the DSP.

DSP run a few complex algorithms: acoustic echo canceling algorithm (AEC), microphone array processing signal algorithm for adaptive beam forming (ABF) and its directivity characteristics, estimation algorithm for direction of arrival (DOA) of useful signal for indoor localization of speaker, in other words speaker room localization, algorithm for reduction of stationary noise, non-steady noise and residual echo (NR- Noise Reduction) and algorithm for system automatic gain control (AGC), because of compensation between different speaker distance from the microphone array. Besides all those basic algorithms, DSP runs some others algorithms more as are: voice activity detector (VAD) on the near-end, VAD on far-end, double talk detector (DTD) on the both sides, additional post filtering (PF) of noise reduction, etc. The aim of mentioned algorithms is maximal reduction of all present noises with minimum of speech signal degradation, therewith secure of transmitting speech signal maximum quality.

Specific aspect of invention subsist adaptive acoustic echo cancellation using an adaptive filter, which mould transferring acoustic way characteristic from loudspeaker to the microphone. Transferring characteristic is complex, working on transmitting way from 2

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS

Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS

Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS

Sync your system to PACER to automate legal marketing.