

平成18年度図書館情報メディア研究科プロジェクト研究 研究成果報告書

種 目	重点配分	共同研究	研究代表者 氏 名	三河 正彦
研究課題	音声・音響・映像情報に基づく環境情報の記録とインデキシング手法に関する研究			
研究組織（研究代表者及び研究分担者）				
氏 名	所属研究機関・部局・職	現在の専門	役割分担	
三河 正彦	図書館情報メディア研究科・准教授	ロボットビジョン	映像情報処理と睡眠覚醒機能に基づく並列処理制御の検討	
田中 和世	図書館情報メディア研究科・教授	音声・音響情報処理	音声・音響情報処理とインデキシングの検討	
研究目的				
<p>近年、センサネットワークに代表されるように、至るところにセンサが設置されている。そこで本研究は、音声・音響・映像センサを用い、常時かつ長期間にわたり音声(話者)を含む環境情報をインデキシングしつつ記録するシステムの構築を目的とする。記録情報にインデックスあるいはアノテーションを付与することにより、情報検索が容易となる。従来は、講演会や映画等、ある特定時間内の音声と映像を対象とした研究が多く、検索対象以外の情報は捨てていた。本研究では、音声と映像に加え音響情報を用いることにより、記録対象の環境情報(環境変化や状況等)について幅広くインデキシング(Rich Indexing)し、さらに多くの付加情報を得ることが可能となる。しかし複数センサを常時かつ長期間用いるには、数多くの情報処理を並列に行う必要があり、また情報処理の要不要や優先順位を状況に応じてその動作を制御し、限られた計算機資源と記録装置を有効利用する必要がある。</p> <p>そこで、次に示す2つの研究課題について研究を進める。</p> <p>(1) 音声・音響・映像データに基づくインデキシング手法の確立：これまでの検索対象である音声と映像情報に加え、データ収集環境に関する情報が含まれる音響データについてもインデキシングすることにより、検索精度の向上と情報の付加が可能となる。例えば音響データには、ドアの開閉等の環境変化や音源の種類や位置等、様々な情報が含まれる。</p> <p>(2) 大容量データの効率のよい蓄積手法の確立：センサ情報はリアルタイム処理が必要なものとそうでないものに分けられる。また長時間データには処理が不要な期間も混在する。人間の睡眠覚醒機能を数理モデル化し、複数センサから得られるデータに対し実行する様々な並列情報処理を動的に制御する手法を提案（研究業績 三河の1, 3, 6, 7及び受賞歴[2]参照）しており、本手法の応用により、必要な情報処理を必要な時に実行、つまり並列に動作する複数の情報処理の動作を動的に制御することができ、効率のよいデータ蓄積が可能となる。</p>				
研究成果				
<p>知覚センサのうち視覚センサを用いた視覚情報処理に関して、複数の視覚センサ、複数の外部情報処理/内部情報処理に適用可能な意識レベルを表現する数理モデル及びセンサ情報を効率よく記憶するための記憶システムを設計・改良し、実機を用いた実験により、その有効性を確認した。これらの研究成果は SICE2006 (International Conference on Instrumentation, Control and Information Technology) や日本ロボット学会、日本知能情報ファジィ学会において発表した。</p> <p>知覚センサのうち音響センサを用いた音環境認識に関して、複数のマイクロホンセンサとする環境雑音の低減処理手法の開発、および複合環境音の分離認識システムの開発を行なった。これらの研究成果は ICSLP2004(International Conference on Spoken Language Processing)や日本音響学会において発表した。</p>				

代表的な研究発表・特許等の成果一覧、特記事項等

<代表的な研究発表>

1. Masahiko Mikawa, Masahiro Yoshikawa, Takeshi Tsujimura, Kazuyo Tanaka : "Intelligent Perceptual Information Processing System Based on Mathematical AIM Model," the 16 th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 2007), 2007 (accepted).
2. Masahiko Mikawa, Masahiro Yoshikawa and Takeshi Tsujimura, "Memory System Controlled by Mathematical AIM Model for Robot Vision Equipped with Sleep and Wake Functions," The SICE-ICASE International Joint Conference 2006, pp. 2272-2277, (2006).
3. Masahiro Yoshikawa, Masahiko Mikawa, and Kazuyo Tanaka, "Real-Time Hand Motion Estimation Using EMG Signals with Support Vector Machines," The SICE-ICASE International Joint Conference 2006, pp. 593-598, (2006).
4. Kazuyo Tanaka, Shi-wook Lee, Yoshiaki Itoh, "Constructing Spoken Document Processing Systems on a Universal Subphonetic Segment Domain," (Invited) Proceeding of 9th Western Pacific Acoustics Conference (WESPAC IX 2006), Paper No.459 (CD-ROM 6 pages), (2006).
5. Kohei Iwata, Yoshiaki Itoh, Kazunori Kojima, Masaaki Ishigame, Kazuyo Tanaka, and Shi-wook Lee, "A Study of Sophisticated Subword Models on a Time Axis for Open-vocabulary Spoken Document Retrieval," Proceeding of 9th Western Pacific Acoustics Conference (WESPAC IX 2006), Paper No.483 (CD-ROM 8 pages) , (2006).
6. Akira Sasou, Futoshi Asano, Satoshi Nakamura, Kazuyo Tanaka, "HMM-based noise-robust feature compensation", International Journal of Speech Communication, vol.48, issue 9, pp.1100-1111, (2006).
7. Kohei Iwata, Yoshiaki Itoh, Kazunori Kojima, Masaaki Ishigame, Kazuyo Tanaka, and Shi-wook Lee, "Open-vocabulary spoke document retrieval based on new subword models and subword phonetic similarity," Proc. of InterSpeech-ICSLP 2006, pp. 325-328, (2006).
8. Yoshiaki Itoh, Takayuki Otake, Kohei Iwata, Kazunori Kojima, Masaaki Ishigame, Kazuyo Tanaka, and Shi-wook Lee, "Two-stage vocabulary-free spoken document retrieval -subword identification and re-recognition of identified sections-", Proc. of InterSpeech-ICSLP 2006, pp. 1161-1164, (2006).