



機能的階層符号化の河川監視への応用

岩橋政宏[†]

[†]長岡技術科学大学工学部電気系 〒940-2188 新潟県長岡市上富岡町 1603-1

あらまし 本報告では、河川の水位監視に特化された新しいタイプの階層符号化（スケーラブル符号化）を提案する。河川映像を撮像するセンサーノードでは、映像信号は時空間基底により帯域分割され、3種類に階層化されたビットストリームとして出力される。第1優先階層には水位検出に有効な帯域信号のみが含まれており、これは水位の常時監視のために超低ビットレートで通信される。水位の「認識」という機能の面で階層化される点が新しい。第2優先階層にはサムネイル映像を閲覧するのに必要な帯域が含まれ、第3の非優先階層には原映像を再生するのに必要な追加データが含まれる。第2、第3の階層については従来の階層符号化と同様である。特に今回は、第1の階層に属するべき帯域を、「認識」の面のみならず「圧縮」の面をも考慮して決定する。結果として、認識性能が3.7%犠牲になったものの、データ伝送量を32.5%削減することが可能となった。

キーワード 水位、河川、認識、スケーラブル符号化

Functionally Layered Coding and its Application to River Monitoring

Masahiro IWAHASHI[†]

[†] Faculty of Engineering, Nagaoka University of Technology, Nagaoka-shi, Niigata, 940-2188 Japan

Abstract In this report, a new type of layered (scalable) video coding especially for the use of water level monitoring is introduced. A sensor node of the system produces a bit stream functionally separated into three layers. The first layer contains the minimum components effective for detecting the water level. It is transmitted at very low bit rate for regular monitoring. The second layer contains signals for thumb-nail video browsing. The third layer contains additional data for decoding the original video signal. A video signal is decomposed into several bands with the three dimensional transform. In this report, optimum bands to be contained into the 1st layer are experimentally investigated considering both of water level detection and data size to be transmitted. As a result, bit rate for transmitting the first layer is reduced by 32.5 % at the cost of negligible 3.7 % decrease of recognition performance for one of video examples.

Keyword Water level, River, Recognition, Scalable coding

1. はじめに

現在、日本ではテレメータと呼ばれる水位監視システムが河川に設置され、観測結果はウェブを介して一般に公開されている[1,2]。しかし、その設置箇所は一級河川などに集中しており、局所的氾濫を起ししやすい小規模用水路などは対象外である場合が多く、ユビキタスな監視に十分な数とは言い難い。これに対し、ウェブカメラであれば比較的lowコストであり、到る所に簡単に設置することができる。しかし、水位を自動的に検出する機能は有していない。

映像から水位を自動的に検出する方法としては、高木らによる方法が知られており実用化もされている[3-5]。この方法では傾斜量水版を河川に設置することで外乱に堅牢な水位検出が可能となる。しかし、板を設置できない場合や板面の汚

れの問題がある。量水板を用いない方法としては、綱島らによる方法が報告されている[6]。これは垂直方向にエッジ検出を行うことで、水平方向の線分として水面を検出する。また、水路壁の苔等の汚れに対しては、フレーム間差分を適用することで、流水の水面のみを検出する。しかし、これにより降雨雪粒子のような動く外乱に対して弱い。

先に我々は、フレームの同期加算を行い流水領域の高周波成分を抑圧した後、水平方向にエッジ検出を行うことで、縦方向エッジの多い陸部分と少ない流水部分に判別し、両者の境界を水位とする、降雨雪粒子にも堅牢な方法を提案した[7,8]。更に、エッジ検出フィルタを2次元に拡張し、その最適設計法も提案した[17-18]。また、テクスチャの分類に使われたガボールフィルタ[11,12]を、JPEG 2000のウェーブレット変換に置換することで、エッジ検出フィルタではなく、基底分解によ

り分割された帯域信号を用いて水位を検出する方法を提案した[9,10]。これらの報告では、流水領域の「認識」に有効な帯域を決定している。

これに対し本報告では、「認識」のみならず「圧縮」をも考慮した際に有効な帯域を決定する。これにより、水位検出という機能の面で階層化された映像符号化（スケーラブル符号化）システムを提案する。河川映像を撮像するセンサーノードでは、映像信号は時空間基底により帯域分割され、3種類に階層化されたビットストリームとして出力される。第1優先階層には水位検出に有効な帯域信号のみが含まれており、これは水位の常時監視のために超低ビットレートで通信される。水位の「認識」という機能の面で階層化される点が新しい。第2優先階層にはサムネイル映像を閲覧するのに必要な帯域が含まれ、第3の非優先階層には原映像を再生するのに必要な追加データが含まれる。これらは従来の SNR ベースの階層符号化[13-16]と変わらない。本報告では特に、第1階層に属するべき帯域を、「認識」の面のみならず、「圧縮」の面をも考慮して決定する。結果として、若干の認識性能が犠牲となるものの、第1優先階層のデータ伝送量が大幅に削減される。なお、本報告の内容の一部は文献[19]にて発表される。

2. 機能的階層符号化

2.1 システムの概要

図1には従来の河川監視システムの一例を示す。映像の「圧縮」処理と水位の「認識」処理の両方を含んでいる。前者は例えば MPEG が汎用的であり、これには動き補償(MC)と離散コサイン変換(DCT)が含まれている。一方、後者にはフレームの同期加算(synchronous frame addition: SFA)とウェーブレット変換(WT)が含まれる[9,10]。一般にセンサーノードは電源確保が困難な河川敷に設置されることが多く、太陽発電と蓄電池で作動するよう、電力消費を出来るだけ抑えたい。

図2は提案法であり、SFA と MC が時間方向のハール変換に統合されている。更に、ウェーブレット変換と DCT が空間方向のハール変換に統合されている。このように「認識」と「圧縮」において3次元直交変換を共有することで、センサーノードにおける処理量を低減して電力消費を抑える。但し、基底としては特にハールに拘るものではないが、構成が簡単のため本報告ではこれを用いた実験結果を示す。

ハール変換は加算器とシフトで構成されるため、WT と MC-DCT の両方を含む従来の構成に比べて簡略化される。なお、図中の Encoding は量子化とハフマン符号化のようなエントロピー符号化を意味する。図1 と比べると図2 では認識処理の一部が受信側で担当される。この点では処理負荷がネットワーク分散されることになる。

デメリットとしては、MC-DCT と比べると符号化効率が劣化する点、時間方向の変換を行うので通信遅延が増加する点、サムネイル映像上の同物体が縺縄化される点が挙げられる。河川監視においては、通常の双方向映像通信の場合と比べ、これらの欠点が運用上あまり問題とならない場合を想定している。

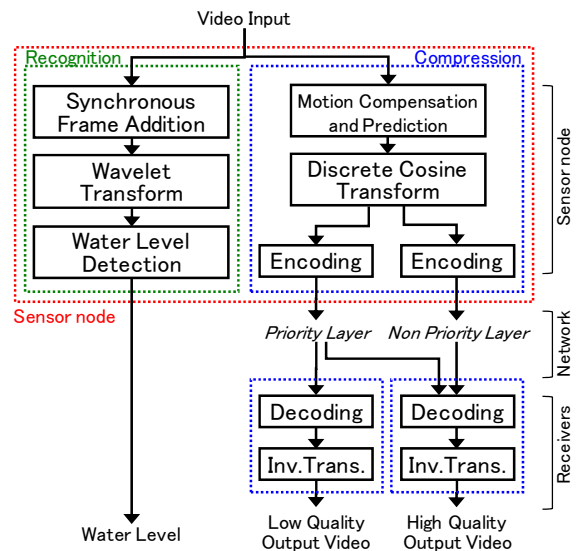


Fig. 1 Conventional tandem system.

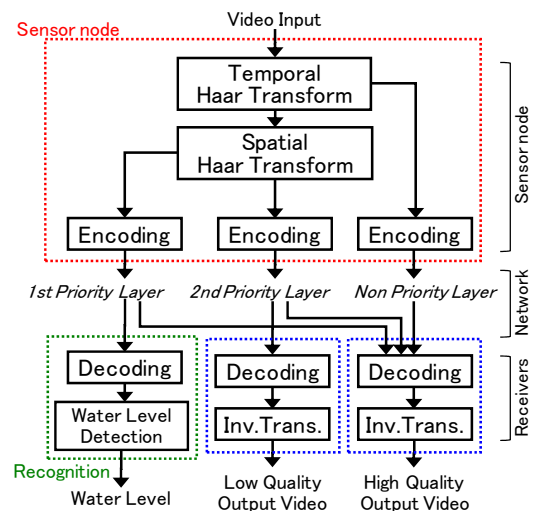


Fig. 2 Proposed unified system.

2.2 水位の検出

河川映像としては、図 4(a)のように、上部に静止した陸部分、下部に流水部分が含まれており、両者の境界を水位として検出する場合を考える。提案法においては、連続する数枚のフレームが同期加算されることで、流水部分の高周波成分が抑圧される。この様子は図 4(b)と図 5(b)により確認できる。提案法の場合、この処理は図 3 に示す時間方向の 2 点ハール変換により実現される。ここで、ハール基底は次式で定義される。

$$\mathbf{H}_{2^{n+1}} = \begin{pmatrix} \mathbf{H}_{2^n} & \mathbf{H}_{2^n} \\ \mathbf{H}_{2^n} & -\mathbf{H}_{2^n} \end{pmatrix}, \quad n=1,2,3,\dots \quad (1)$$

$$\mathbf{H}_2 = \begin{pmatrix} 2^{-1/2} & 0 \\ 0 & 2^{-1/2} \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

最も低い周波数の帯域成分は、SFA を適用した結果と本質的に等しい。提案システムでは、この帯域を第 1 および第 2 優先帯域とする。

更に、空間方向の 2 点ハール変換を縦と横に施すことで、第 1 ステージにおける 4 つの帯域信号{1LL, 1HL, 1LH, 1HH}を生成する。とくに最低域の LL 帯域については、更に変換を施して第 2 ステージにおける{2LL, 2HL, 2LH, 2HH}を生成する。この処理を第 S ステージまで繰り返すことで信号をオクターブ分割する(今回は $S=5$) [9,10]。

任意のステージにおける LL 帯域は第 2 優先階層に分類する。これは従来のスケラブル符号化と同様である。第 1 優先階層には、水位の検出に必要な最小限な帯域を含める。これについては第 3 節に述べる。

図 4(c)には流水領域に属すると判別された画素を黒、陸領域を白で示している。白い画素の数を各行ごとに数えることでその行の特徴量とする[7]。結果を図 4(d)に例示する。この特徴量の大小により領域を二分する点を水位とする。

2.3 最尤推定 (ML estimation)

各画素は、1 つあるいは複数の帯域信号を特徴ベクトルとして、陸 ($k=0$) あるいは流水 ($k=1$) のいずれかに判別される。この際、帯域信号は原画像と同じ空間解像度に拡大される。判別に用いる帯域信号の数が、その画素に対する特徴ベクトルの次元である。ここで、位置 (n_1, n_2) における画素の特徴ベクトル $\mathbf{G}(n_1, n_2)$ を、ガウス型の確率密度関数:

$$P(\mathbf{G}(n_1, n_2)|T_k) = \frac{1}{2\pi\sqrt{|\mathbf{C}_k|}} \exp\left(-\frac{d_k^2(n_1, n_2)}{2}\right), \quad (2)$$

として近似する。このとき、各画素のクラス T_k ($k=0$ or 1) を Mahalanobis 距離:

$$d_k^2(n_1, n_2) = [\mathbf{G}(n_1, n_2) - m_k]^T \mathbf{C}_k^{-1} [\mathbf{G}(n_1, n_2) - m_k] \quad (3)$$

により判別する。ここで、 m_k および \mathbf{C}_k は、クラス T_k における特徴ベクトルの重心および共分散行列をそれぞれ表す[9-12]。これらは、あらかじめ指定された各クラスの教師領域により与えられる。

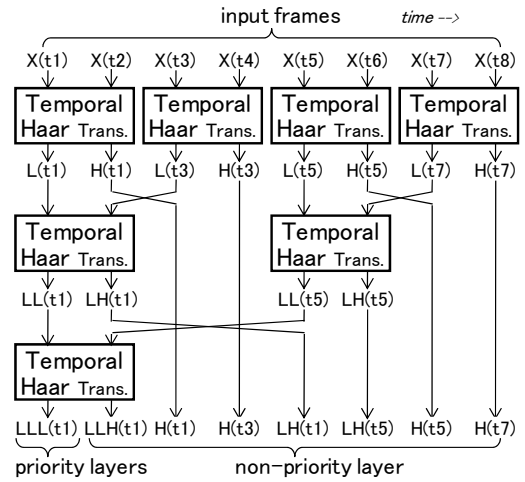


Fig. 3 Octave decomposition with the temporal HT.

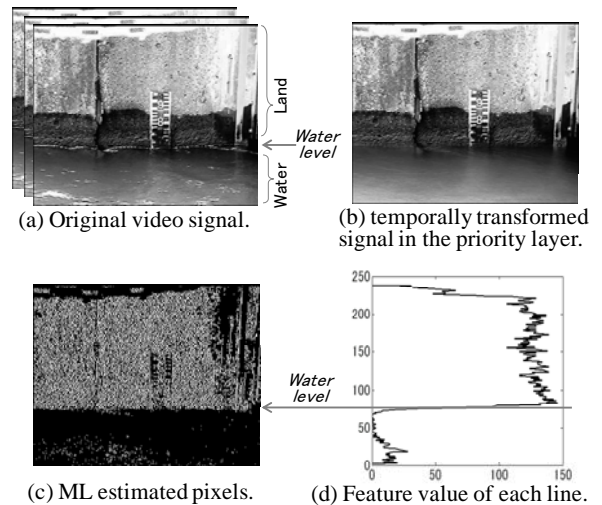


Fig. 4 Signal processing of the proposed system.

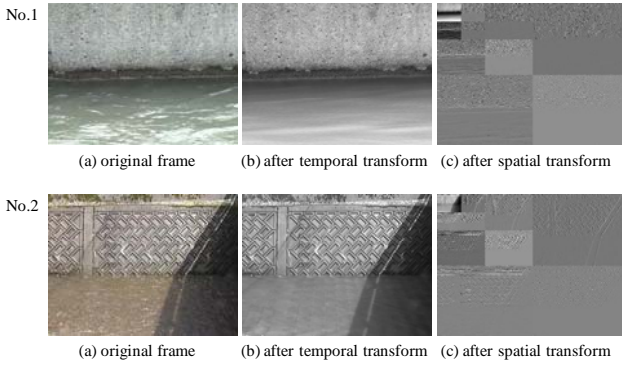


Fig. 5 Tested video samples [320x240 px1, 30 fps].

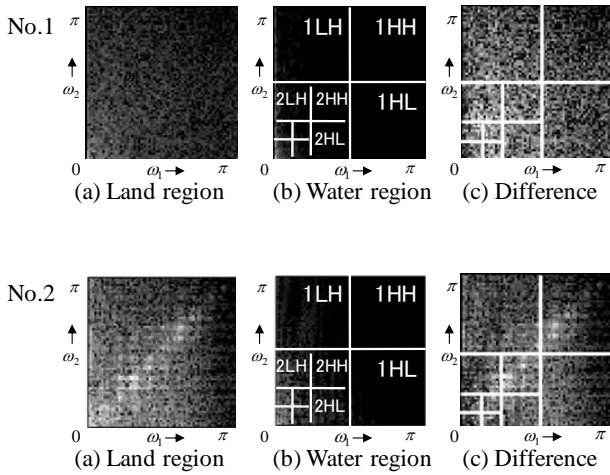


Fig. 6 Spectrum of the video samples.

3. 実験結果

3.1 1次元の特徴ベクトルを用いた場合

既に述べたように、時間方向のハール変換により得られた最低域成分は、SFAの結果と等価であり、流水領域における高域成分が抑圧される[7-10]。この様子は信号の周波数（対数）振幅スペクトルを示す図6からも確認できる。水位の検出に必要な帯域は、流水領域と陸領域のスペクトルの相違（対数では差）が存在する帯域と考えられる。図5の映像サンプルに対する実験結果を、図7-8および表1-2にまとめる。

映像サンプル no.1 の場合について以下に述べる。図7(a)の縦軸は、画素毎に流水か陸かを判別する際の誤り確率（以下、誤判別率: discrimination error rate “ R ”）である。横軸は、その際に1次元特徴ベクトルとして用いた1つの帯域である。2HHを用いた時は $R=4.57$ [%]であり、認識の点では最も良い。また、2HLでは $R=4.94$ であり2番目である。これらの帯域は図6(b)に示されるように中周波である。1LLや2LLは認識においては効果がなかった。スペクトルと認識率との関係については今後更なる検討が必要である。

図7(b)には、水位の検出誤差（単位はピクセル）と誤判別率との関係を示す。両者の相互相関係数は0.853であった。一般に判別誤差が少なければ水位検出誤差も小さくなると考えられるが、両者は必ずしも線形な関係にあるとは言えないようである。

本報告では、「認識」のみならず「圧縮」をも考慮して第1優先階層に含めるべき帯域を決定する。そこで、誤判別率に対するもう一つの尺度として、特徴ベクトルとして用いる帯域信号のデータサイズ（data size “ B ”）を導入する。図7(c)に実験結果を示す。3HHを特徴ベクトルとした場合、 $B=3.2$ [KB]となり、データ伝送量（圧縮）の面では最良である。本報告では、 R と B の両方を勘案する。すなわち、次式で定義される標準化距離(normalized distance “ ND ”)を用いる。

$$ND = \sqrt{\frac{NR^2 + NB^2}{2}}, \quad (4)$$

$$NX = 50 + 10 \cdot \frac{X - m_X}{\sigma_X}, \quad X \in \{R, B\}.$$

この新しい規準により、3HLが第1優先階層として決定されることになる。これは、図7(c)において原点に最も近い位置にある。

文献[9,10]における先の研究報告（以下、従来法）では、「認識」のみを考慮して特徴ベクトルとなるべき帯域が決定されている。これにより、認識の点で最も良い2HHが第1優先階層となる。これは、図7(c)において最も低い位置にある。結果として、2HHが従来法、3HLが提案法となる。

サンプル映像 no.2 に対する結果は図8に示す。この場合は、2HHが従来法、2HHが提案法となり、両者の違いは無い。

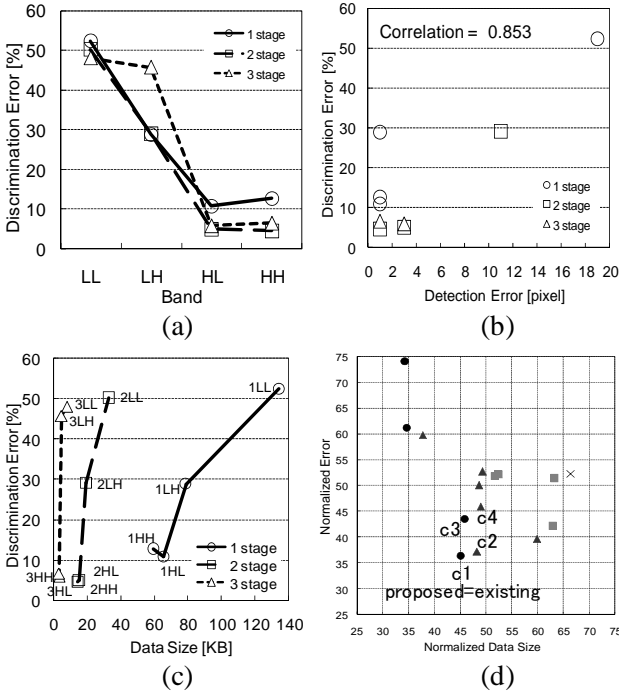


Fig. 7 Experimental results for sample no.1.

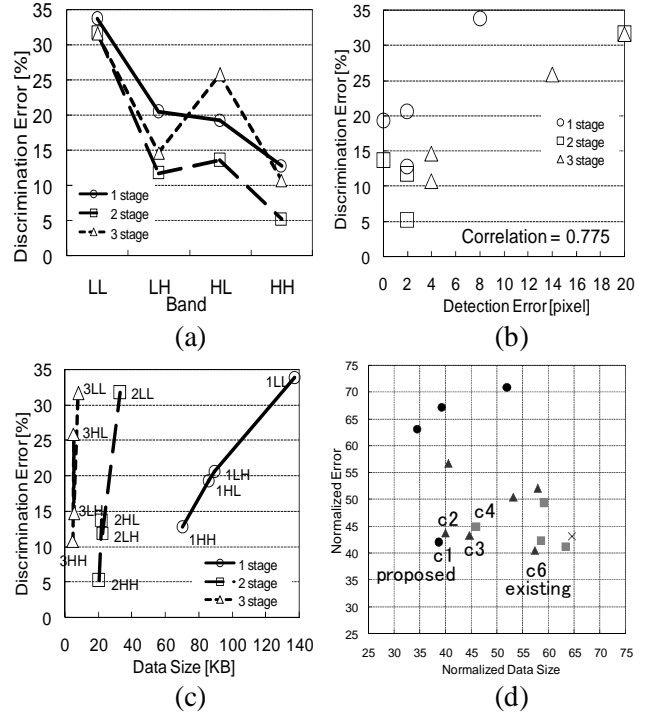


Fig. 8 Experimental results for sample no.2.

Table 1: Combination of bands for the 1st priority layer.

sample 1				
#	band	normalized data size	normalized error	normalized distance
		NB	NR	ND
	<i>stdev</i>	10.00	10.00	5.30
	<i>mean</i>	50.00	50.00	50.67
	<i>min</i>	34.30	36.36	41.01
	<i>max</i>	66.44	74.04	59.75
c1	2HH	45.18	36.36	41.01
c2	3HH-2HH	48.28	37.16	43.08
c3	2HL	45.90	43.44	44.69
c4	2HL-3HH	48.99	45.93	47.49
c5	3HL-2HH	48.66	50.02	49.35
c6	3HL	34.69	61.18	49.73
c7	3HL-3HH	37.78	59.79	50.01
c8	2HL-2HH	59.87	39.65	50.78
c9	3HL-2HL	49.37	52.71	51.07
c10	3HL-3HH-2HH	51.76	51.81	51.78
c11	3HL-2HL-3HH	52.47	52.11	52.29
c12	2HL-3HH-2HH	62.97	42.15	53.58
c13	3HL-2HL-2HH	63.35	51.42	57.69
c14	3HH	34.30	74.04	57.70
c15	3HL-2HL-3HH-2HH	66.44	52.21	59.75

Table 2: Combination of bands for the 1st priority layer.

sample 2				
#	band	normalized data size	normalized error	normalized distance
		NB	NR	ND
	<i>stdev</i>	10.00	10.00	5.79
	<i>mean</i>	50.00	50.00	50.62
	<i>min</i>	34.54	40.45	40.36
	<i>max</i>	64.63	70.81	62.10
c1	2HH	38.70	41.95	40.36
c2	3HH-2HH	39.95	43.71	41.87
c3	2LH-2HH	44.70	43.25	43.98
c4	2LH-3HH-2HH	45.96	44.85	45.41
c5	2LH-3HH	40.54	56.71	49.29
c6	2HH-1HH	57.37	40.45	49.64
c7	3HH	34.54	63.03	50.82
c8	3HH-2HH-1HH	58.62	42.25	51.10
c9	3HH-1HH	53.21	50.35	51.80
c10	2LH-2HH-1HH	63.37	41.15	53.43
c11	2LH-3HH-1HH	59.21	49.29	54.48
c12	2LH-3HH-2HH-1HH	64.63	43.05	54.91
c13	2LH	39.29	67.13	55.00
c14	2LH-1HH	57.96	52.01	55.06
c15	1HH	51.95	70.81	62.10

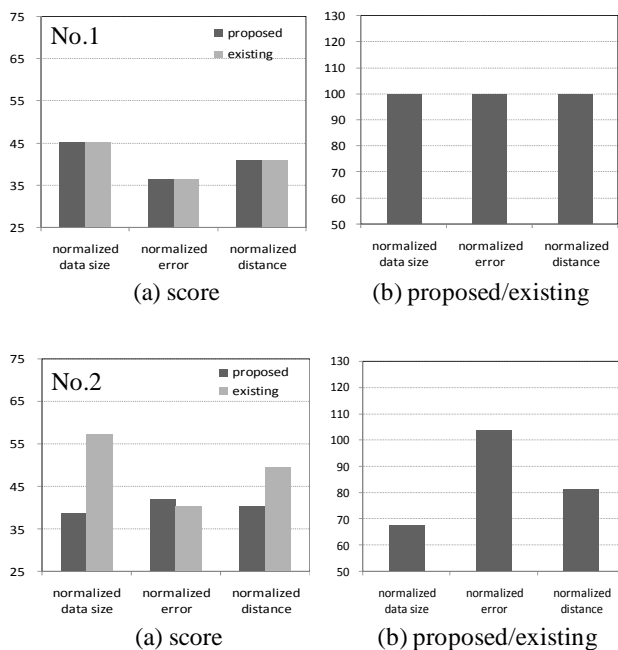


Fig. 9 Effectiveness of the proposed method.

3.2 多次元の特徴ベクトルを用いた場合

複数の帯域を組み合わせて多次元の特徴ベクトルとして判別に用いた場合の結果を、図 7(d)、図 8(d)と表 1, 2 に示す。1 次元特徴ベクトルとして認識率の面でベスト 4 を選び、これらの全ての組み合わせ (${}^4C_4 + {}^4C_3 + {}^4C_2 + {}^4C_1 = 15$ 通り) について調査した。

サンプル映像 no.1 の場合、表 1 に示すように、従来法では NR に従って 2HH が選ばれる。提案法では ND に従って 2HH が選ばれる。この場合、両者の違いは無い。図 7(d)は横軸を NB 、縦軸を NR としてプロットしてある。従来法は最も低い位置、提案法は最も原点に近い位置にある。

サンプル映像 no.2 の場合、表 2 に示すように、従来法では NR に従って組合せ番号 c6 の 2 次元特徴ベクトル 2HH-1HH が選ばれる。一方、提案法では ND に従って組合せ番号 c1 の 1 次元特徴ベクトル 2HH が選ばれる。

従来法と提案法について、 $\{NB, NR, ND\}$ を比較した結果を図 9 に示す。サンプル映像 no.1 の場合は、図 9(a)に示されるように、両者の差は無い。一方、サンプル映像 no.2 の場合には、図 9(b)に示されるように、第 1 優先階層の伝送に必要なデータ量を示す NB は 32.5 % 削減される。但し、誤判別率 NR は 3.7 % 増加している。これは、「圧縮」を考慮したために「認識」性能が若干犠牲になっていることを意味する。しかし、図 8(b)に示されるように、この程度の変化では水位検出誤差にはほとんど影響がなく、この意味では無視できるほど小さい。

4. まとめ

河川監視に特化された機能的階層符号化を提案した。送信側であるビデオセンサノードからは水位検出という機能の面で階層化されたビットストリームが出力される。第 2 優先階層及び非優先階層は従来の SNR ベースのスケラブル符号化と同様であるが、第 1 優先階層は水位検出に有効な帯域が選ばれる。この際、認識と圧縮の両方を考慮した標準化距離(normalized distance: ND)が選定規準として用いられる。ある映像サンプルの例では、認識のみを考慮した従来法に比べ、提案法ではデータ伝送量が 32.5 % 削減されることが示された。

謝辞

本研究の一部は、総務省・略的情報通信研究開発推進制度(SCOPE-C) 一情報通信分野における戦略的な競争的研究資金一、“地域防災のための人と環境にやさしいインテリジェント・ビデオセンサー・ネットワークに関する研究”により実施されている。

文 献

- [1] Ministry of international affairs and communications, Japan, <http://www.soumu.go.jp/english/index.html>
- [2] Ministry of Land, Infrastructure and Transport Government of Japan, <http://www.river.go.jp/>
- [3] Y. Takagi, A. Tsujikawa, M. Takato, T. Saito, M. Kaida, "Development of a non-contact liquid level measuring system using image processing", 7th International workshop on Instrumentation, control and automation of water and waste water treatment and transport systems, pp.409-415, 1996.
- [4] Y. Takagi, T. Yoneoka, H. Mori, M. Yoda, A. Tsujikawa, T. Saito, "Development of a water level measuring system using image processing", The 1st IWA Conference on Instrumentation, control and automation, pp.309-316, 2001.
- [5] Y. Takagi, A. Tsujikawa, M. Takato, T. Saito, M. Kaida, "Development of a non-contact liquid level measuring system using image processing", Water science and technology, vol. 37, no.12, pp.381-387, 1998.
- [6] N. Tsunashima, M. Shiohara, S. Sasaki, J. Tanahashi, "Water level measurement using image processing", Information processing society of Japan, Research report, Computer vision and image media, vol.121, no.15, pp.111-117, March 2000.
- [7] A. Saito, M. Iwahashi, "Water Level Detection Algorithm based on Synchronous Frame Addition and Filtering", IEICE circuits and systems Karuizawa workshop, pp.525-530, April, 2006.
- [8] M. Iwahashi, S. Udomsiri, "Optimum Filter Design for Water Level Detection", IEICE Tech. report, SIP2007-3, IE2007-3, ISSN 0913-5685, vol.107, no.22, pp.13-18, April 2007.
- [9] Y. Imai, M. Iwahashi, "Water level detection based on wavelet transform of JPEG2000", IEICE circuits and systems Karuizawa workshop, pp.531-534, April, 2006.
- [10] M. Iwahashi, S. Udomsiri, Y. Imai, S. Fukuma, "Water Level Detection for River Surveillance utilizing JP2K Wavelet Transform", IEEE Asia Pacific Conference on Circuits and Systems (APCCAS), pp.1766-1769, Oct. 2006.
- [11] D. Dunn, W. E. Higgins, "Optimal Gabor filter for texture segmentation", IEEE Trans. image processing, vol.4, no.7, pp.947-964, 1995.
- [12] H. Tanaka, Y. Yoshida, K. Fukami, H. Nakano, "Texture segmentation using amplitude and phase information of Gabor filters", IEICE Trans. D-II, vol.J84 -DII, no.12, pp.2565-2576, Dec. 2001.
- [13] M. Iwahashi, W. Kameyama, K. Ohyama, N. Kambayashi, "A Motion Compensation Technique for Down-Scaled Pictures in Layered Coding", IEICE Trans. on Communications, vol.E-77, no.8, pp.1007-1012, Aug. 1994.
- [14] B. R. Lee, K. K. Park, J. J. Hwang, "H.263-based SNR scalable video codec", IEEE Trans. Consumer Electronics, Vol.43, Issue 3, pp. 614-622, Aug. 1997.
- [15] Li Weiping, "Overview of fine granularity scalability in MPEG-4 video standard", IEEE Trans. on circuits and systems for video technology, vol. 11, issue 3, pp. 301 - 317, March 2001.
- [16] A. R. Reibman, L. Bottou, A. Basso, "Scalable video coding with managed drift", IEEE Trans. Circuits and Systems for Video Technology, Vol.13, Issue 2, pp.131-140, Feb. 2003.
- [17] H. Kikuchi, Y. Wada, A. Saito, M. Iwahashi, "Optimum Design of Temporal-Spatial Filter for Detection of Running Water Region in Video", IEICE Circuits and Systems Karuizawa Workshop, pp.203-208, April, 2007.
- [18] M. Iwahashi, "Water Level Detection from Video with FIR filtering", Sixteenth International Conference on Computer Communications and Networks (ICCCN), Aug. 2007. (to be published)
- [19] M. Iwahashi, "Water Level Detection for Functionally Layered Video Coding", IEEE International Conference on Image Processing (ICIP), Nov. 2007. (to be published)