

情報発生量の分布に基づく遠隔講義撮影の自動化

大西 正輝[†] 泉 正夫[†] 福永 邦雄[†]

Automatic Production of Video Images for Distance Learning System
Based on Distributed Information

Masaki ONISHI[†], Masao IZUMI[†], and Kunio FUKUNAGA[†]

あらまし 近年、ネットワーク技術の発展に伴い、映像情報などの大容量データ転送が可能となり、大学や各種の学校ではサテライトシステムなどを用いた遠隔講義が行われている。講義風景を撮影する方法として、1台の固定したカメラを用いる方法や、複数台のカメラをディレクタなどが各時点でスイッチングする方法が考えられる。しかし、前者では講義の臨場感が伝わらず、また板書が見にくいといった問題点があり、後者ではカメラを制御する人手が必要となる。本論文では、動画像処理により講義映像の撮影・編集を自動化する手法を提案する。まず、映像中に発生した画素ごとの情報量を定義する。次に、その情報の分布に基づいて複数台のカメラを制御し、あらかじめ生成したスイッチング規則に従い映像の選択を行うことで、受講者に必要な情報を多く含む臨場感のある講義映像をリアルタイムで生成する。また、受講者の視認テストにより本手法の有効性を確認した。

キーワード 遠隔講義, 情報発生量, 臨場感, 動画像処理, カメラ制御

1. ま え が き

近年、ネットワーク技術の発展に伴い、大容量映像データの転送が可能となり、大学や予備校をはじめとした各種の学校では通信衛星などを用いた遠隔講義が盛んに取り入れられている [1]~[3]。遠隔講義は、文字通り空間的に離れた講義者と受講者が仮想クラスルームの中で一同に会して学習を行うことである。アメリカでは、衛星通信によって配信される教育プログラムにより、大学院修士の学位を取得することが可能になっている。また日本でも、平成 10 年度からは大学間で一定の条件を満たした遠隔講義についての単位認定が認められ、遠隔講義の教育効果などについても盛んに研究されている。

遠隔講義を実現するにあたり、講義風景を撮影する簡便な手法として 1 台の固定したカメラにより撮影した映像を転送する手法が広く用いられているが、講義の臨場感が伝わらず、「講義の雰囲気」、「画像の見やすさ」といった観点から問題が生じる [3]。この問題を回避する一つの手法として、複数のカメラマンが講義の

状況に応じたカメラワークで撮影し、それらの映像の中からディレクタが各時点で最適な映像に順次スイッチングするといった手法が取り入れられている。この場合には、固定カメラに比べて受講者に必要な情報を多く含む臨場感のある講義映像の撮影が可能となるが、カメラマンやディレクタといったカメラを制御する人手が必要となる。また、これらの人手を削減しようとする講義者自らが映像取得のための作業を行わねばならない事態が生じる問題がある。

そこで、映像取得及びその支援法に関係する研究が行われている [4]~[9]。プレゼンテーション支援法として教材呈示装置などに利用することを目的としたズームパラメータの決定手法 [4] や、講義を自動的に記録編集して教材化するもの [5] などが挙げられる。また、テレビ会議システムの映像生成を支援する手法としてテレビ番組のカメラワークの知識を用いて画面を切替える手法 [6] では、固定カメラによる会議システムとの比較を行い高い評価を得ている。

講義映像の撮影支援法としては、複数のカメラで取得した映像から、コンテキストをどれほど適切に撮影できているかを採点することでスイッチングを行う手法 [7], [8] や講義者の位置・姿勢の情報からカメラワークを決定する手法 [9] などが挙げられるが、講義の雰

[†] 大阪府立大学工学部, 堺市
College of Engineering, Osaka Prefecture University 1-1,
Gakuen-cho, Sakai-shi, 599-8531 Japan

困気を伝えるまでには至っておらず、人手による撮影には及ばないところが多い。

本論文では講義者を中心とした講義映像を撮影・編集するにあたり、一般にカメラマンの仕事である撮影カメラでのトラッキング及びズーム、またディレクタの仕事であるカメラのスイッチングを画像処理によって抽出した情報発生量の分布に基づいて行い、臨場感のある講義映像を自動的に撮影・編集する手法を提案する [10]。まず、時系列の講義映像から空間的に存在する情報と時間的に存在する情報に着目し、これらをもとに情報発生量を定義する。次に、講義シーンを対象とした映像から算出した情報発生量の分布をもとに情報の多い領域を中心にパン・チルトカメラを制御し、用意した生成規則に従いスイッチングを行うことで臨場感のある映像をリアルタイムで自動的に生成する。また、受講者の視認テストにより本手法の有効性を確認している。

2. 情報発生量

2.1 色情報の確率密度

まず、注目する画素が近傍領域の画素と比べて、どの程度類似した色情報を有しているかを確率密度関数を用いて調べる。

自然風景の画像中の近傍画素間には強い相関があるため、近傍領域における画素の色情報分布は多変数正規分布に従うと仮定する。このような仮定のもとでは、近傍画素と類似する色情報をもつ画素の発生する確率は高くなり、近傍画素と異なる色情報をもつ画素の発生する確率は低くなる。そこで、注目する画素位置 $P_i = [x_i \ y_i]^T$ がその近傍領域の色情報分布からみて、どの程度の確率で発生するかを確率密度関数の形で与える。確率密度関数は、画素 P_i の色情報 $X_i = [R_i \ G_i \ B_i]^T$ を用いて次式の 3 次元多変数正規分布に従うと仮定する。

$$p(X_i) = \frac{1}{(2\pi)^{\frac{3}{2}} |\Sigma|^{\frac{1}{2}}} \exp \left\{ -\frac{1}{2} (X_i - \mu)^T \Sigma^{-1} (X_i - \mu) \right\} \quad (1)$$

ここで、 μ 、 Σ は近傍領域における n 画素の色情報の平均値及び共分散行列を表し、

$$\mu = \frac{1}{n} \sum_{k=1}^n X_k \quad (2)$$

$$\Sigma = \frac{1}{n} \sum_{k=1}^n (X_k - \mu)(X_k - \mu)^T \quad (3)$$

で与える。確率 $p(X_i)$ は、近傍領域の色情報を考慮した確率密度関数であり、画素 P_i が色情報 X_i を発生する確率を表していると考えられる。

2.2 空間情報量と時間情報量

次に画像中のそれぞれの画素が持っている情報量を定義する。

Shannon の情報理論を用いて、画像を解析する研究はこれまでも行われている [11], [12]。Jägersand は、解像度に関する情報量を定義することで、最適な解像度を知る手法を提案している [11]。また、竹内らは輝度値を情報量の尺度とし、それを用いて複雑な部分を選択的に注視するアクティブビジョンシステムを開発している [12]。本手法においても、Shannon の情報量を用いて統一的に特徴量を扱う。

本論文では、映像中の各々の画素が情報をもつと考え、画素 X_i のもつ情報量を式 (1) を用いた次式で定義する。

$$H(X_i) = -\log_2 p(X_i) \quad (4)$$

これは、確率密度関数に Shannon の情報量の考え方を適用したもので、発生する確率の小さい画素が発生した場合ほど、その画素から大きな情報が得られるとしている。画像中のすべての画素について式 (4) を計算することで、画像上の情報の分布を調べることができる。

ここでは、2 種類の情報について取り上げる。空間的な色情報の変化による情報と時間的な色情報の変化による情報である。空間的な変化による情報とは、時間軸の標準化によって得られる各々の 2 次元画像（静止画像）における情報であり、絵柄が複雑な領域ほど大きな情報量をもつ。一方、時間的な変化による情報とは、時間軸標準である 2 次元画像の相互相関による情報であり、画像中の動きの変化が大きい領域ほど大きな情報量をもつ。

今、注目する画素 X_i を中心として、図 1 のように時間的に連続して得られる入力画像平面上に $m \times n$ 画素、時間方向に l 画素の近傍領域を考えた場合、画像平面上 $m \times n$ 画素の近傍領域を考えた画素 X_i の情報量を空間情報量 $H_s(X_i)$ 、時間軸上 l 画素の近傍領域を考えた画素 X_i の情報量を時間情報量 $H_t(X_i)$ と呼ぶことにする。

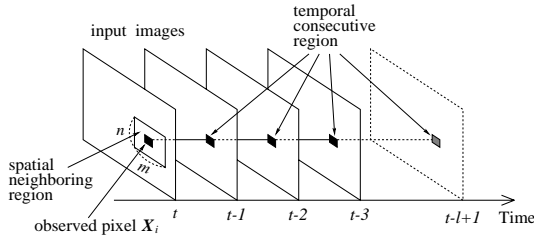


図1 近傍領域

Fig.1 Neighboring and consecutive region.

図2に入力画像(a)から得られた空間情報量(b), 時間情報量(c)を濃淡画像により示す. それぞれ, 濃淡値の濃い領域ほど情報量大きいことを示している. 空間情報量の大きい領域は講義者や文字などの領域であり, 時間情報量の大きい領域は現在身振りを行っている講義者の領域であることが確認できる.

2.3 空間・時間の相互情報量

講義映像を撮影する際には, 受講者にとって必要な情報を多く含む領域を撮影する必要がある. 受講者にとって必要な情報とは, 空間的・時間的の両方に情報をもつ部分である. 例えば, 講義者の身振り手振りであり, 黒板に発生した文字である. 一方, 空間的な情報があったとしても, 以前に書かれた文字は情報を失っている. また, 時間的な情報があったとしても, 消された部分には受講者の注目する情報は存在しない. もちろん, 空間と時間に独立に発生するランダムノイズにも注目すべき情報は存在しない.

そこで, 受講者の注目する情報を式(4)で定義した時間情報量と空間情報量に相互情報量の考え方を適用した次式で定義する.

$$\begin{aligned}
 I(X_i) &= \log_2 \frac{p_{st}(X_i)}{p_s(X_i) \cdot p_t(X_i)} \\
 &= H_s(X_i) + H_t(X_i) - H_{st}(X_i)
 \end{aligned}
 \tag{5}$$

ここで, $H_{st}(X_i)$ は時空間情報量であり, 画像平面上 $m \times n$ 画素と時間軸上 l 画素の近傍 $l \times m \times n$ 画素の領域を考えた画素 X_i の情報量を表す.

式(5)は, 画素 X_i における時間情報量と空間情報量の共通に含まれる情報量と解釈することができ, 情報発生量とする. つまり, 本論文では注目画素 X_i における空間的な変化はその画素を中心とした現在のフレームに代表して現れ, 時間的な変化はその画素位置における時間的な変化に代表して現れるものと考えて情報発生量を定義している. 情報発生量 $I(X_i)$ は, 時



(a) 入力画像
(a) Input video image.



(b) 空間情報量
(b) Amount of spatial information.



(c) 時間情報量
(c) Amount of temporal information.



(d) 時空間情報量
(d) Amount of spatiotemporal information.



(e) 情報発生量
(e) Amount of generated information.



(f) 情報発生量(2値画像)
(f) Amount of generated information (binary image).

図2 情報発生量

Fig.2 Amount of generated information.

間と空間に共通の情報が発生した場合に大きな値を持ち, 時空間の情報量を用いることで, 時間と空間に独立に出現するノイズ情報を軽減する特徴をもつ. 図2に入力画像(a)から抽出した時空間情報量(d), 情報発生量(e)を濃淡画像により示す. また(f)には, 情報発生量をしきい値処理して2値画像にしたものを示す. 現在情報を持っている身振りを行っている講義者のみを抽出できていることが確認できる. 以後の処理

では、計算の高速化を図るため情報発生量をしきい値処理により 2 値化したものを取り扱う。

3. 情報発生量の分布に基づくカメラ制御

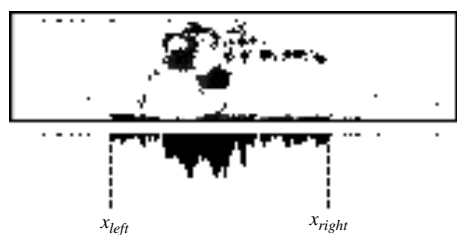
2.3 で定義した情報発生量 $I(X_i)$ は、画素 X_i に発生した情報量を抽出するものであり、同じ場所にある続ける情報は、抽出の対象にならない。しかし現実には講義シーンを考えると、黒板に書かれた文字などの情報は、書かれた直後に情報を失うわけではない。受講者が板書する時間や、理解する時間など個人差があるものの一定の時間、情報は残っているものと考えられる。そこで空間情報量の高い画素については、2 値化した情報発生量が一定時間残るものとする。図 3 (b) に情報が 15 秒程度残るとしたときの情報発生量の分布を示す。ここでは、現在書いている文字だけではなく、ある程度の過去に書かれた文字情報も抽出できていることがわかる。

受講者が講義映像を見る場合には、多くの情報が含まれている領域を注視する。そこで固定カメラより得られた入力画像から抽出した情報発生量の多い領域へパン操作でカメラを制御し、受講者に必要な情報を多く含む講義映像を生成する。

まず、しきい値処理により 2 値化した情報発生量 (空間情報量の高い領域では、一定時間残るものとする) に対して水平方向の射影ヒストグラムを生成する (図 3 (b))。このヒストグラムには、画面の左端



(a) 入力画像
(a) Input video image.



(b) 情報発生量とそのヒストグラム
(b) Amount of generated information and its histogram.

図 3 情報発生量とそのヒストグラム

Fig. 3 Amount of generated information and its histogram.

を原点として右方向を正とする座標系を設定する。次にノイズ除去を施した後、一定以上の大きさをもつ連続したヒストグラムの左端の座標 x_{left} 及び右端の座標 x_{right} を求める。そして、カメラの視線方向は $(x_{right} + x_{left})/2$ に、ズームは $(x_{right} - x_{left})^{-1}$ に比例するようにカメラアングルを設定することで、より多くの情報が撮影できるようにカメラを制御する。このようなカメラ制御により、カメラの視線は確実に情報の中心方向をとらえ、情報が広い範囲に分布している場合にはシーン全体を見渡せるようにズームアウトで撮影し、反対に情報が狭い範囲に集中している場合には、その領域をズームインして撮影する。

しかし、間断なくカメラを動かした場合には映像として見づらくなることが考えられるため、ここでは撮影カメラから講義者を見失いそうになった時点で、カメラアングルを変更し情報を追跡する。具体的には、視線方向 $(x_{right} + x_{left})/2$ が現在のカメラの視線よりしきい値以上にずれた場合にカメラアングルを変更する。

4. 複数カメラのスイッチング

3. では、1 台のカメラでの撮影を考えた場合に、より多くの情報を含む撮影法について述べた。しかし、1 台のカメラでの撮影では講義者自身によって黒板の文字が隠される場合が生じる。また、撮影中にカメラ位置を動かせないとすると一定方向からのみの撮影映像となるため、映像の変化に乏しくなり受講者が飽きてしまうという問題がある。ここでは、複数の位置にカメラを設置した場合に、撮影カメラを効果的に選択するスイッチングの問題について検討する。

講義における適切な映像を選択する基準としては、「黒板の文字が良く見える」、「講義者と視線があっている」などが考えられる。また、黒板の右端を教室の左端に設置したカメラから撮影すると黒板が斜めになるために板書にひずみを生じ違和感を感じる。

これらのことを考慮し実際に講義映像の撮影・編集を試み、経験的にカメラのスイッチング規則を作成した。スイッチングが度重なると受講者は戸惑いを感じることや、カメラ台数が多くなるとスイッチング規則が複雑になること、また撮影カメラ台数を一定以上に増やしても台数に比例した効果が期待できないことなどを考慮して、ここでは撮影用の制御カメラを 2 台とした。それぞれのカメラにおいてあらかじめ黒板の両端の座標を調べておき、講義者は黒板から大きく離れ

ないという仮定のもとで、認識用カメラで求めた撮影方向から座標変換を施すことでそれぞれのカメラ位置での撮影方向を計算した。実験に用いた講義室の大きさは約 650 [cm] 四方である。図 4 のようにカメラを設置した場合におけるスイッチング規則を表 1 に示す。上段の条件ほど優先順位が高いものとする。また、図 4 での 1 目盛は 50 [cm] 四方を表している。

ズームアウトして撮影する場合には、講義室の広範囲が映るため、教室の端から撮影した場合には黒板などにひずみを生じる可能性がある。そこで、撮影カメラは中央の *CameraI* とした。黒板の右端付近に情報が分布している場合には中央に設置したカメラを、黒板の左端付近に情報が分布している場合には左に設置したカメラを用いて撮影する。これは、ともに映像にひずみを生じないようにするためである。また、 x_{right} に変化がなく、 x_{left} が大きく左に動いている場合には、講義者が情報の左側にいるものと考えられ

る。この場合に左側に設置した *CameraII* で撮影すると、講義者によるオクルージョンが生じ、映像中の文字領域に死角が増えると考えられるため、反対方向の *CameraI* で撮影する。反対に x_{left} に変化がなく、 x_{right} が大きく右に動いている場合には *CameraII* で撮影する。また、これらの条件を満たさない場合には、スイッチングは行わずに撮影を続ける。初期の撮影カメラは、*CameraI* とした。

一方、「講義者と目線があっている」に関する問題は映像中に発生した情報をとらえる本手法では解決することが困難であるため、講義者に撮影カメラがわかるように配慮した。

5. 実験及び考察

5.1 実験

固定カメラから得られた講義映像から情報発生量を算出し、その結果をもとに図 4 のように配置した 2 台のパン・チルト及びズーム制御が可能なカメラを制御して、リアルタイムで講義映像を撮影・編集した。認識用に使用した画像サイズは 640 × 160 画素で取り込んだ画像の解像度を粗くして 160 × 40 画素にしたものであり、黒板の両端が映像中に収まるように画像処理用カメラを配置した。1 台の PC を用いて画像処理（情報発生量の計算）と 2 台のカメラ制御及びスイッチングの処理を行った。また、処理速度は 1 秒に 2 フレーム程度である。情報発生量は空間情報量の高い画素については 15 秒程度残るものとした。論文中の l, m, n は統計での標本の大きさを表し、ある程度の大きさをもつことが望ましい。しかし、値が大きくなると処理に時間がかかるため、時間内に十分な処理ができない。本実験では $l = m = n = 3$ とした。本手法による撮影映像の例を図 5 に示す。

5.2 アンケート評価

本手法の有効性を確認するために受講者による視認テストをアンケート評価によって行った。本手法についての説明を行った 20 分程度の講義映像を撮影し、大学生の受講者 35 人に評価してもらった。評価の対象となる映像は次の 3 種類である。

- 固定カメラを用いて撮影した映像
- 人手により撮影した映像
- 本手法を用いて自動で撮影した映像

人手による撮影はカメラマン 2 人とディレクタ 1 人で行った。また、カメラの配置は、図 4 と同じであり、自動撮影と同じ条件にした。ともに専門家ではないが、

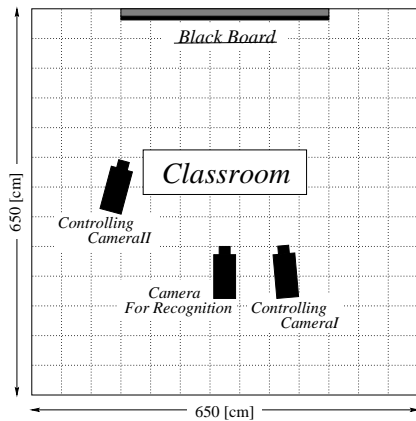


図 4 カメラ配置

Fig. 4 Camera layout in the classroom.

表 1 カメラのスイッチング規則
Table 1 Rules of camera switching.

CameraIで撮影	
条件	説明
$(x_{right} - x_{left})^{-1}$: 十分小さい	板書のひずみを考慮
$(x_{right} + x_{left})/2$: 黒板右端付近	板書のひずみを考慮
x_{right} : 変化なし \cap	板書の隠れを考慮
x_{left} : 大きく左に変化	
CameraIIで撮影	
条件	説明
$(x_{right} + x_{left})/2$: 黒板左端付近	板書のひずみを考慮
x_{left} : 変化なし \cap	板書の隠れを考慮
x_{right} : 大きく右に変化	



図5 本手法による撮影映像例
Fig.5 Video images made by this method.

あらかじめ数日間に撮影の練習を行った。撮影手法の採点方法は、まず認識用のカメラ位置に設置した1台の固定カメラにより撮影した映像を受講者に見てもらい、その評価を3として人手による撮影映像と自動撮影による映像をそれぞれ5段階で点数をつけてもらった。また、受講者に撮影方法は知らせておらず、アンケートは複数回行い、呈示順を変えることで誤差が相殺されるように注意した。

アンケート項目は、文献[2]及び[6]を参考にして本手法に関係すると思われる5項目を選んだ。それぞれの集計結果を表2に示す。人手・自動の各項目の値は、それぞれの評価の平均点を表す。また、表中F値は撮影手法間の平均平方と誤差の平均平方の比を表しており、値が大きいほど撮影手法間に評価の違いが現れていることになる。

このアンケート結果をもとに分散分析法を用いて、人手によって撮影した映像と本手法によって自動撮影した映像に差がないという仮説の検定を行った。映像間に差がないという仮説は、すべての項目でF値が $F(1, 68; 0.05) = 3.98$ よりも小さいため、有意水準5%の検定で棄却することができない。つまり、人手の撮影による講義映像と本手法による自動撮影の講義映像の間に、はっきりとした差を認めることはできないと結論できる。特にQ.4において、F値が大きいのは、講義者や板書の文字から情報量を抽出することができているため、それぞれの領域を追跡することはできているが、本手法では講義者や板書などといった情報源を特定することができないため、受講者が真に見たい点を決定することができないことに起因してい

表2 アンケート結果
Table 2 Results of questionnaires.

質問項目	人手	自動	F 値
Q.1 講義者の表情はよくわかる	4.36	4.23	0.58
Q.2 板書(図表)はよくわかる	4.47	4.23	2.85
Q.3 講義者のジェスチャはよくわかる	3.86	3.80	0.06
Q.4 自分の見たい点がよく見える	4.23	3.91	3.23
Q.5 講義の雰囲気伝わります	3.87	3.72	0.51

● 固定カメラ映像の評価値 3.00 に対する評価

ると考えられる。

また、本手法による撮影映像と人手による撮影映像において違和感を感じるカメラワークがあった時間及び、見たいところを見られなかった時間を受講者に計ってもらった。受講者は10人であり、評価対象となる映像は前記のアンケートに用いた映像と同じ講義映像である。本手法による自動撮影によって生成された映像では、それらの合計時間が全体の講義時間のうちの平均11.45%であったのに対し、人手による撮影での時間は平均6.67%であった。これらの結果が表2の結果に反映されているものと考えられる。一方、自動撮影では平均2.29分に1回の割合で受講者が違和感を感じる映像があったのに対して、人手による撮影では、2.43分に1回の割合であった。人手による撮影と自動撮影で違和感を感じた回数は、ほとんど変わらないが人手による撮影の方がその時間が短いのは、カメラマンが映像中の違和感を察知しカメラアングルをすぐに変更できるのに対して、本手法の自動撮影では発生した情報が一定時間残るために、カメラアングルの変更にかかるためと考えられる。

5.3 考 察

本手法について考察するとともに、実験・アンケートにより生じた問題点を調べる。

- Shannon の情報理論により、空間的かつ時間的に発生した情報を情報発生量として定義した。これは、映像中に発生した情報を理論的に意味づけられたものと解釈できる。また、本手法で定義した情報発生量は時空間の情報をを用いることで、ノイズの影響を除去した情報を抽出することができる。

- 映像から抽出した情報発生量の分布をもとにカメラワークを決定することで、臨場感のある講義映像を自動で撮影する手法を提案した。また、映像を撮影するのに最適なカメラ位置を決定するスイッチング規則を作成した。本手法による自動撮影は、注目点をアップで撮影することが可能になるため受講者に必要な情報を多く含む映像生成が期待できる。

- 本手法による自動撮影映像を評価するため、アンケートを行った。受講者によるアンケートの回答結果から、本手法は従来の固定カメラのシステムに比べ、講義者の表情、板書、ジェスチャがわかりやすいなどの評価が得られた。また、人手による撮影までには至らなかったが、すべての質問項目において有意に差を開けられることはなかった。

- 本手法では、発生した情報を中心に撮影を行っているが、画像処理やカメラ制御などの時間遅れにより実際の講義者の位置と撮影位置にずれを生じる場合がある。この問題を回避するには、講義者の行動を予測する必要がある。また、カメラ制御の誤差による撮影のずれも存在するため、得られた撮影画像からカメラ制御の誤差を補正するなどの手法が必要となる。これらは、認識カメラが 1 台である本手法のみでの解決は困難であり、今後の課題である。

- 本手法で作成したスイッチング規則は、抽出した情報発生量をもとに制御カメラが 2 台であることを想定した場合でのスイッチング規則である。ここでは、中央・左・右の撮影位置で映像の見やすさを考えているため、本手法を単純に撮影カメラが 4 台以上の場合に拡張することは困難である。一般的に、カメラ台数を増やした場合におけるスイッチング規則を生成するには、更に多くの情報を抽出する必要があると考えられる。例えば、板書・講義者といった情報源を抽出することができれば、カメラ台数を増やした場合におけるスイッチング規則生成への対応も容易になり、カメラワークにいつそう専門家の知識を取り入れることが

できると考えられる。この問題は今後の課題である。

6. む す び

本論文では、入力動画像から空間的な情報と時間的な情報を Shannon の情報理論に基づいて定義し、それら二つに共通した情報量を情報発生量とした。また、情報発生量の分布をもとにカメラを制御し、経験的に作成したスイッチング規則を用いて、臨場感のある講義映像を自動で撮影する手法を提案した。最後に、本手法により撮影した講義映像を受講者による視認アンケートで評価することで有効性を確認した。

謝辞 撮影規則を構築する方法をアドバイスして頂いた毎日放送(株)の指川正情報システム部長に心から感謝致します。また、たび重なるアンケートに快く応じて下さった本学情報工学科の学生諸君に深く感謝致します。

文 献

- [1] S. C. Brofferio, "A university distance lesson system: Experiments, services, and future developments," IEEE Trans. on Education, vol.41, no.1, pp.17-24, Feb. 1998.
- [2] 宇井 修, 中山 実, 清水康敬, "衛星通信講座における講義形態と学習者評価の関係," 信学論 (D-II), vol.J80-D-II, no.4, pp.892-899, April 1997.
- [3] 清水康敬, 前迫孝憲, "キャンパス間を結ぶテレビ講義の評価," 信学論 (A), vol.J69-A, no.10, pp.1181-1189, Oct. 1986.
- [4] 戸田真志, 川嶋稔夫, 青木由直, "画像情報分布に基づく視覚ズームの適応的制御," 信学論 (D-II), vol.J81-D-II, no.1, pp.84-92, Jan. 1998.
- [5] S.X.Ju, M.J.Black, S. Minneman, and D. Kimber, "Analysis of gesture and action in technical talks for video indexing," Proc. of Computer Vision and Pattern Recognition, pp.595-601, June 1997.
- [6] 井上智雄, 岡田謙一, 松下 温, "テレビ番組のカメラワークに基づいた TV 会議システム," 情処学論, vol.37, no.11, pp.2095-2104, Nov. 1996.
- [7] 美濃導彦, 亀田能成, "講義状況の取得方式と取得映像の教育応用に関する考察," クラスルームシンポジウム講演論文集, pp.43-50, Jan. 1998.
- [8] 宮崎英明, 吉吉健太郎, 亀田能成, 美濃導彦, "複数カメラを用いた講義映像の実時間作成法," MIRU'98, pp.I-123-I-128, July 1998.
- [9] 大西正輝, 松本昌紀, 福永邦雄, "動画像処理による遠隔講義映像の自動生成とその評価," 信学技報, ET98-62, Oct. 1998.
- [10] 大西正輝, 泉正夫, 福永邦雄, "情報発生量による遠隔講義映像の自動生成とその評価," 信学技報, PRMU98-176, Jan. 1999.
- [11] M. Jägersand, "Saliency maps and attention selection in scale and spatial coordinates: An information

theoretic approach,” Proc. of 5th Int. Conf. on Computer Vision, pp.195-202, June 1995.

- [12] 竹内義則, 大西 昇, 杉江 昇, “情報理論に基づいたアクティブビジョン,” 信学論 (D-II), vol.J81-D-II, no.2, pp.323-330, Feb. 1998 .

(平成 11 年 2 月 24 日受付, 6 月 21 日再受付)

大西 正輝 (学生員)

平 9 阪府大・工・情報卒. 平 11 同大大学院博士前期課程了. 現在同大大学院博士後期課程在学中. ロボットビジョンに関する研究に従事. 電気学会, 日本ロボット学会, 映像情報メディア学会各学生員.

泉 正夫 (正員)

昭 58 阪府大・工・電気卒. 昭 60 同大大学院博士前期課程了. 同年シャープ(株)入社. 平 2 より阪府大・工・電気助手. 現在同情報助教授. 画像認識・理解に関する研究に従事. 情報処理学会, システム制御情報学会, IEEE 各会員. 工博.

福永 邦雄 (正員)

昭 42 阪府大・工・電気卒. 昭 44 同大大学院修士課程了. 同年同大・工・電気助手. 現在同大・工・情報教授. コンピュータビジョン, グラフ理論とその応用などの研究に従事. 情報処理学会, 人工知能学会, システム制御情報学会, IEEE 各会員. 工博.