

# LN-7 チョーク音による板書動作の開始と終了の判定 A Judgment of Beginning and Ending of Teacher's Chalking with Sounds

芦川 平† 菅沼 明‡ 谷口 倫一郎‡  
Taira Ashikawa Akira Suganuma Rin-ichiro Taniguchi

## 1. はじめに

現在、様々な教育機関において遠隔講義が行われている。その講義の撮影方法は、カメラを固定したまま行うか、または人間がカメラを操作して行うかの2つに分けられる。しかし、前者は一定領域だけを撮影するため十分な情報が得られず、また後者はコストがかかるという問題がある。そこで、我々の研究室では、黒板を使う講義を自動的に効率よく撮影するシステム(Automatic Camera control system for Education 以下 ACE)を構築している[1][2]。ACEでは、教師が新しく板書した領域を一定時間ズームして撮影するという手法をとっている。その際に必要になるのが、教師が板書を開始したと終了したことの判定である。現在のACEでは画像処理のみを用いてこの判定を行っているが、教師の位置によっては板書が隠れてしまう等の原因により失敗することがある。そこで、本研究では板書動作時に生じるチョークの音を利用し、板書動作の判定を行った。

## 2. ACEについて

### 2.1 板書領域の抽出

教師は新しく板書した領域(以下、最新板書領域)を説明対象とする可能性が高い。ACEでは、最新板書領域を一定時間ズームして撮影し、その映像を学生側に提供する。新しい文字が認識できた場合に、教師による板書の開始と判定し、一定時間新しい文字が認識できなければ、板書の終了と判定する。この判定により、最新板書領域の抽出を行う。

### 2.2 問題点

画像情報のみで板書領域を抽出する場合、板書動作の開始と終了の誤判定により、以下の場合にうまく領域が抽出できないことがある。

- 新しい文字の隠蔽(教師が立ち止まって板書した場合等)
  - 黒板の移動(移動式黒板の場合)
- これらの場合、画像情報のみで板書動作の開始と終了の判定をするのは難しい。そこで、講義中に発生するチョーク音に注目し、板書動作の判定を行う。

## 3. チョーク音による板書動作の判定

### 3.1 講義中に発生する音

黒板を用いた講義において、講義中に発生する音としては教師や学生の声、チョーク音が考えられる。その他にも様々な音が発生することもあるが、それらはノイズとして考えることができる。

### 3.2 チョーク音の特徴

チョーク音には以下の特徴が見られる。

- 衝突音である
- 音の発生間隔が一定時間以上ひらく

チョーク音はチョークと黒板の衝突によって起こる瞬間音である。また、人間の動作には限界があるので、その間隔はある一定値以上は必ず開くことになる。板書中のチョーク音を調べたところその間隔は図1のようになった。

図2に、チョーク音の波形とフレーム毎のスペクトル分布を示す。20msをフレーム間隔としてスペクトル分布を調べると、チョーク音があるフレームにおいて3kHz付近のスペクトルが急増する。さらに、チョーク音の終了とともに次のフレームで急増する。この3kHz付近の周波数帯を特徴周波数帯と呼び、教室、黒板、チョークの種類によって多少異なる。図3に、チョーク音の特徴周波数帯におけるフレーム間のスペクトル和の隣接差分を示す。このグラフにより、チョーク音が発生すると特徴周波数帯のスペクトル和の差分が突出し、さらに次のフレームでは、負の方向に突出することがわかる。

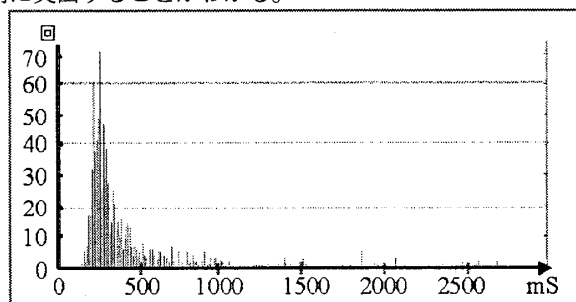


図1 チョーク音の時間間隔

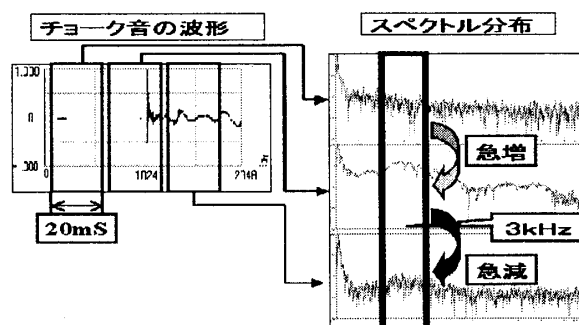


図2 チョーク音の波形とスペクトル分布

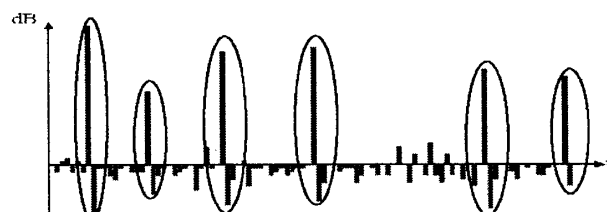


図3 隣接フレームのスペクトル和の差分

†九州大学大学院システム情報科学府

‡九州大学大学院システム情報科学研究院

### 3.3 チョーク音の認識

チョーク音の重要な特徴は衝突音である。これは、音声には見られない特徴であり、また講義中に連続して発生する衝突音は少ない。そこで、チョーク音の認識は、衝突音の検出により行う。

#### 3.3.1 前処理

まずチョーク音のみが複数録音されたデータを用意する。すべてのチョーク音を周波数分析し、特徴の表れる周波数帯 B を求める。その後、以下の手順でスペクトル和の差分の突出量の閾値(T1,T2)を求める。

- (1) Flag=False; stack1 と stack2 を空にする
  - (2) 録音された音データをフレームに分割し、各フレーム毎の B のスペクトル和 I(n)を計算する
  - (3) 隣接差分  $\Delta=I(n)-I(n-1)$  を求め
    - (ア) Flag=False かつ  $\Delta > TH$  ならば push(stack1,  $\Delta$ ); Flag=True;
    - (イ) Flag=True かつ  $\Delta > TH$  ならば pop(stack1); push(stack1,  $\Delta$ );
    - (ウ) Flag=True かつ  $\Delta \geq 0$  ならば pop(stack1); Flag=False;
    - (エ) Flag=True かつ  $\Delta < 0$  ならば push(stack2,  $\Delta$ ); Flag=False;
  - (4) (2)(3)をすべてのデータに対して繰り返す
  - (5) stack1, stack2 に積まれた値の平均をそれぞれ求め  $\alpha$  倍したものを T1, T2 とする
- ただし、TH は閾値、 $\alpha=0.7$  とする。

#### 3.3.2 認識処理

実際にチョーク音の認識を行う際には、取得したデータに対してリアルタイムで以下の処理を行う。

- (1) Flag=False;
- (2) B のスペクトル和 I(n)を計算する
- (3) 隣接差分  $\Delta=I(n)-I(n-1)$  を求め
  - (ア) Flag=False かつ  $\Delta > T1$  ならば Flag=True;
  - (イ) Flag=True かつ  $\Delta < T2$  ならば チョーク音; Flag=False;
  - (ウ) Flag=True かつ  $\Delta \geq T2$  ならば Flag=False;
- (4) (2)(3)を繰り返す

#### 3.4 評価

チョーク音の認識の評価をするために、ある文章を書く板書動作の音を録音したものを使用して実験した。使用したのは、チョーク音のみの音(音ファイル 1)とチョーク音と音声とを混合した音(音ファイル 2)の 2 つである。これらの音ファイルに対して認識率を求めた結果を表 1 に示す。ただし理想回数とは、実際にチョーク音がする回数を文字の画数から求めた。

表 1 チョーク音の認識率

音ファイル	理想回数	認識回数	認識率(%)
1	433	303	70
2	433	278	64.2

ファイル 2 で認識率が下がるのは、チョークの音より、教師の声が大きくチョークの音が消される場合があるからである。

#### 3.5 板書動作の判定

板書動作中はチョーク音が連続して発生していると考えられる。板書動作中のチョーク音の時間間隔は図 1 より、最頻値が 250ms、平均が 380ms、標準偏差が 280ms、最大

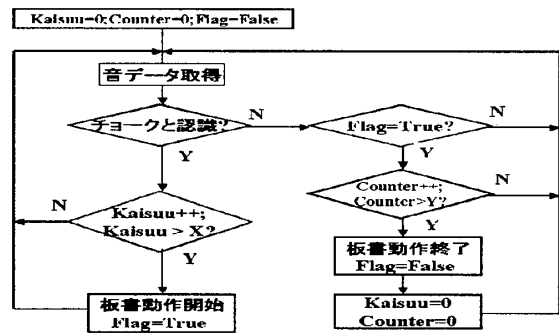


図 4 チョーク音による板書動作の判定

値が 2960ms であることがわかる。そこで、1 秒間にチョーク音を 3 回以上認識したときを板書開始とし、その後 3 秒間チョーク音がとぎれたときを板書終了とした。このアルゴリズムを図 4 に示す。

## 4. 実験と評価

本手法を評価するために、模擬講義を行いチョーク音による板書動作の判定を行った。人による板書動作の判定と、チョーク音による板書動作の判定を比較した。この結果を表 2 に示す。ただし、時間は講義を始めてからの経過時間である。

板書中にもかかわらず、板書動作終了と判定している場合がある。これは、教師が板書中に教科書を参照するなどして、3 秒以上板書を中断した場合であった。しかし、画像情報による板書動作の判定では失敗する板書が隠れた場合でも、チョーク音の認識による板書動作の判定が正確に行えることを確認した。

表 2 板書動作の開始と終了の判定

人による板書動作の判定		チョーク音による板書動作の判定	
開始時間	終了時間	開始時間	終了時間
1:50	2:45	1:53	2:44
3:30	4:25	3:34	4:21
5:00	6:31	5:04	5:35
		5:40	6:38
7:15	8:23	7:21	8:25
9:10	10:20	9:13	10:15
10:50	12:13	10:54	11:56
		12:00	12:16

## 5. おわりに

今回の研究では、講義中のチョーク音の認識により、板書動作の判定を行った。ACE においては、画像情報と音情報の両方を合わせることで、板書動作の判定精度をあげることができると考えている。

#### 参考文献

[1] 錦織修一郎, 菅沼明, 谷口倫一郎, “黒板講義を対象とした講義自動撮影システム”, 電子情報通信学会, 信学技法, Vol.100, No.701, pp.79-86, 2001.

[2] A.Suganuma and S.Nishigori, “Automatic Camera Control System for a Distant Lecture with Videoing a Normal Classroom”, Proc. World Conf. on Educational Multimedia, Hypermedia and Telecommunications, to appear, 2002.