唇周辺画像からの唇抽出方式の検討

1L－3

古山 浩志、井上 郁夫
松下電器産業(株)東京通信システム研究所

1. はじめに

音声認識において、音声と共に話者の口唇領域を含む映像を併用することにより雑音環境下における認識率の向上が報告されている[1]。そこで，我々はこの性質に着目し、音声認識技術を映像検索に利用することを目的として、視覚情報と聴覚情報を統合して音声認識の精度を向上するための方式開発を行っている。

読唇方式に関しては、話者の唇周辺に貼ったマークから抽出した特徴量を用いて認識実験を進めている[2]。今回は話者の唇周辺領域を含む画像から、読唇に用いる特徴量（唇の上下左右４点）を自動抽出するための方式の基礎検討として、色差成分を用いた唇抽出を行ったので報告する。

2. 色差成分ヒストグラムによる唇抽出

色成分を用いた唇抽出方式として、RGB 成分の R/G 値の色ヒストグラムに対して閾値を設定することにより唇抽出をする報告などがある[3]。

今回は唇を抽出するための色差成分として(R-G-B)を用い、色差成分ヒストグラムから唇抽出のための閾値を話者の個人差に依存せずに自動抽出する方式を検討した。

A Study of Lip Extraction Method from Lip Region Image.
H. Furuyama and I. Inoue.
Matsushita Electric Industrial Co., LTD.
4-5-15, Higashi-Shinagawa, Shinagawa-ku, Tokyo, 140-8632, Japan.

図１は対象とする話者（女性、口紅あり）の唇周辺領域を切り出した画像の例である。

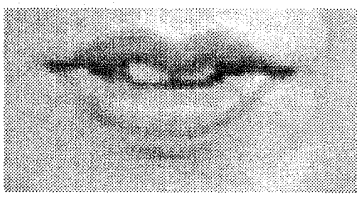

図１．唇周辺画像

図２に唇周辺部分の R-G-B 成分ヒストグラムを示す。図では、ヒストグラムに２つのピークがあり、R-G-B 成分が-10 付近のピークは唇部分によるもの、-60 付近のピークは唇周辺の肌色部分によるものである。

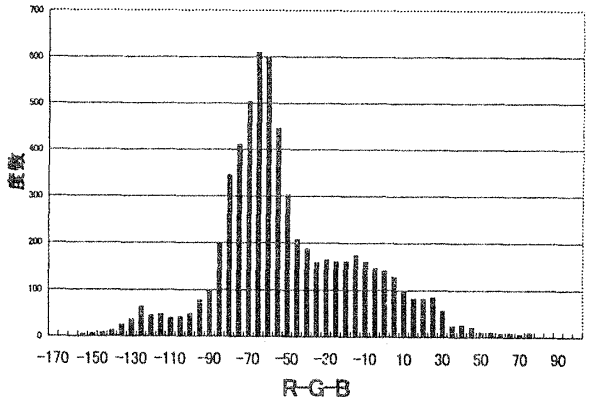

図２．色差成分（R-G-B）ヒストグラム

R-G-B 成分ヒストグラムから、唇部分と唇周辺の肌色部分の分離を行うため、図３に示すように、A 領域：唇部分と比較的広い範囲で唇周辺の肌色部分を含む領域、B 領域：ほぼ唇部分のみを含む領域（A 領域と B 領域の面積比は３：１とした）を切りだし、A 領域、B 領域の R-G-B 成分ヒストグラム、A 領域と B 領域のヒストグラムの差分のグラフを作成した（図４）。

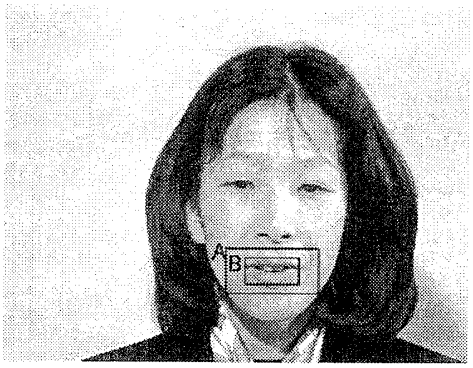

図３．ヒストグラム作成領域（領域A,B）

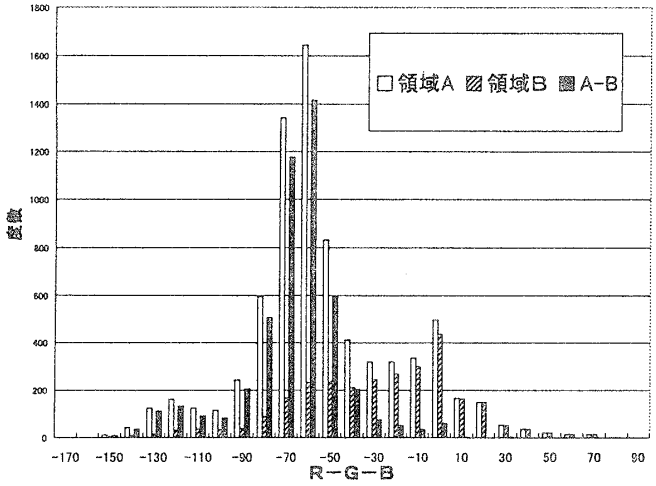

図４．色差成分（R-G-B）ヒストグラム（領域A,B,A-B）

図のヒストグラム（A-B）は、（唇部分を含まない）唇周辺の肌色部分のR-G-B成分ヒストグラムとなっており、この唇周辺の肌色部分ヒストグラムから唇部分を抽出するための閾値を設定することが可能である。

図５は、唇周辺の肌色ヒストグラムから設定した閾値を用いて２値化することにより唇を抽出した画像である。また、図６には、唇領域の上下左右の４点（読唇のための特徴点）を自動抽出し、４点にマークを表示した画像を示す。

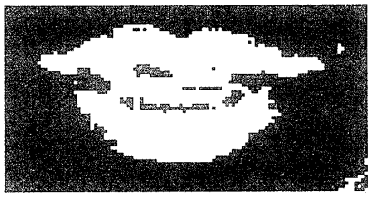

図５．唇抽出画像

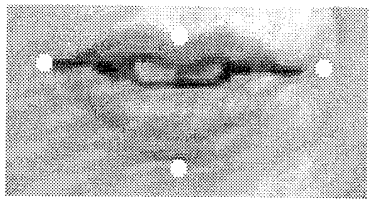

図６．特徴点抽出画像

１０名の話者（男性６名、女性４名）に対して本手法による唇抽出実験を行い、唇部分を良好に抽出できることを確認した。

3．まとめ

読唇に用いる特徴量自動抽出の基礎検討として、唇領域を含む話者の画像から唇部分の抽出のため色差成分ヒストグラムによる唇抽出実験を行った。

色差（R-G-B）成分から唇部分を抽出するための閾値を設定する方法として、唇を含み異なる領域の２つの画像の色成分ヒストグラムの差分から唇抽出のための閾値を設定する方式を提案、検証した。

謝辞：本研究は、通信・放送機構からの委託研究「インテリジェント映像技術の研究開発」の一環として行ったものである。

参考文献：

1）Shintani et al., “An Isolated Word Speech Recognition Using Fusion of Auditory and Visual Information.”, IEICE Trans. Fundam. Electoron., Vol.E-79-A No.6 pp777(1996).

2）古山他、「唇周辺のマークを用いた音素認識のための基礎検討」、情報処理学会第５７回全国大会、2-37(1998).

3）岡田他、「唇の形状抽出の一方法」、電子情報通信学会春期全国大会、D-178、pp.7-160(1989).