修　士　論　文

題　　目

論文題目

指導教員

松永　昭一 教授

平成29年度

長崎大学大学院 工学研究科

総合工学専攻

野崎　大智 (52117321)

目 次

図目次

表目次

第1章

研究背景

1.1 研究背景

近年、通信・放送業界では地上デジタル放送の開始や、新たな高速通信規格の誕生など、通信ネットワークの急速な発達が見られる。それに伴い、誰もがテレビやパソコンだけでなくスマートフォン・タブレットなど様々なデバイスを通して手軽に膨大な量の音声・映像データを入手し、好きな時に好きな場所で視聴することが容易な時代となった。しかし、入手できる情報量が増えた分、それら全てが必要であるとは限らず、自分に必要な情報のみを手軽に取捨選択できれば便利である。映像・音声データに、話者や内容のインデックスの情報が付与されていれば、その部分だけを選択して視聴できる。しかし、世の中には膨大な量の映像・音声データが存在するため、それら全てに人手でインデックスを付与することは事実上不可能である。そこで、自動的にインデクシングすることが望まれる。

自動でインデクシングを行うためには、映像・音声データ内の発話区間、発話者、発話内容の特定が必要である。これらを推定する技術のことをダイアライゼーションと呼び、本研究はこの技術の実現を目指す。

本研究では、世の中に存在する映像・音声データの中である特定の人物に情報が集中する形式で行われるニュース番組に着目した。ニュース番組は主にアンカー (司会役のアナウンサー) を中心として、アンカーがレポーターなどに話を振りながら、ニュースが進行していく。また、ニュース番組は収録環境が良いため、研究対象としても適しており、ニュース番組で高精度にダイアライゼーションができると、同じスタイルのその他の映像・音声データにも用いることができると考えられる。そこで、本研究ではニュース番組を対象として研究を行う。

1.2 研究目的

ダイアライゼーションで推定する情報の中で、本研究では発話者の特定に着目した。本研究で対象とするニュース番組には以下の特徴がある。

ニュース番組の特徴

- 30 分程度のニュース番組の中で複数の多様な話題がある
- 1 人または複数のアンカーおよび天気予報士など複数の話者が存在する
- 話者情報（話者数、性別、話者の声質など）および発話区間が未知である

このようなニュース番組において、ニュースの話題にインデキシングが行われていることは必要な話題の検索に重要である。ニュース番組のアンカーには以下のような特徴があり、インデキシングには重要な情報を持つと考えられる。

アンカーの特徴

1. 発話数が多い
2. ニュース番組の司会および話題の切り替えを行う
3. ニュース番組の全体にわたって発話している

このため、アンカーの発話区間のみの音声認識を行うことによって、より高精度なインデキシングが実現可能であると考えた。これまでに先行研究として音響特徴毎に発話者群を二分化していき、段階的に分類していく手法によって話者識別を行う研究が行われた [1]。しかし、この先行研究では話者の発話区間は既知であるとして行われたため人手によって発話区間を切り出す必要があった。そこで、本研究では、話者情報と発話区間が未知の場合でも用いることが可能なアンカーの発話区間抽出手法を提案する。

1.3　論文構成

次章以降における本論文の構成は、まず 2 章で音声認識システムの概要について説明を行う。次に 3 章では話者識別システムの概要として i-vector、コサイン類似度の理論的背景の説明、および使用する音声データの概要の説明を行い、ニュース番組音声における発話間の i-vector のコサイン類似度を用いることによる効果を検証する。4 章では i-vector を用いた単純なアンカーの発話区間抽出アルゴリズムによる発話区間抽出実験を行い、問題提起を行う。5 章では本研究で提案するアルゴリズムの説明を行う。6 章では提案手法を用いた発話区間抽出実験を行い、本研究における提案手法を用いることで話者情報と発話区間が未知の場合におけるアンカーの発話区間抽出への効果を検証する。7 章では 6 章で抽出したアンカーの発話区間の音声認識を行い、どれほど単語が正確に認識されているかを検証する。 8 章では本研究において検証された実験の結果を元に結論を述べる。

第 2 章

理論的背景

2.1 音源識別

音響データ（ニュース音声、会話音声等）の中にはさまざまな音源種別（声、音楽、雑音等）の音が混在している。音源識別とは、音響データ中に含まれる音源種別を自動的に識別することである。ここでの処理は、音響データのスペクトル解析を行い、音響特徴パラメータを求め、あらかじめ用意した各音源種別の音響特徴パラメータの分布と比較することで音源種別を識別する。

本研究では、ニュース番組の音声データに音響特徴パラメータを用いた音源識別 [9] を用い、音声データ中の音源種別を以下の 4 つに分類した。

(1) 音声区間: アナウンサーやインタビューの声

(2) 音楽区間: オープニングやエンディングなどの音楽、BGM

(3) 背景雑音区間: 自動車走行音や鳥の泣き声

(4) 無音区間: 音量が極めて小さい区間

また、音源識別システムは音響データを各種別へ識別するための音響特徴パラメータの分布に混合ガウス分布を用いている。本研究では、混合数 8 のガウス分布を使用している。本研究の音響特徴パラメータを表 2.2、各音源種別の識別性能 F 値を表 2.3 に示す。ただし、表の評価値とは、音源種別間のデータ数を考慮するために、各音源種別の F 値に各音源種別の割合をかけ、合計したものである。

表 2.2: 音源識別のための音響特徴パラメータ

以下に、音源識別のためのスペクトル解析と音響特徴パラメータについて説明する。

Ftinoyatu

2.1.1 スペクトル解析

音響データのスペクトル解析の手法として最も一般的に利用されている方法は、短時間フーリエスペクトル分析がある。この方法は、音響データから連続する数 10ms 程度の時間長の信号区間を切り出し、切り出された信号が定常性（一定周期で繰り返す）と仮定して、スペクトル解析を行う。

スペクトル解析の流れは以下の通りである。

(1) フレーム化処理:与えられた信号 $s(n)$ に長さ N の分析窓を掛けることで以下のような信号系列 $s_\omega(m;l)$ を取り出す。

$$S_w(m;l) = \sum_{m=0}^{N-1} \omega(m)s(l+m) \qquad (l = 0, T, 2T, \cdots) \tag{2.1}$$

ここで、添え字 l は信号の切り出し位置に対応している。すなわち、l を一定間隔 T で増加させることで定常とみなされる長さ N の信号系列 $s_\omega(n) \quad (n = 0, 1, \cdots, N-1)$ が間隔 T で得られる。この処理をフレーム化処理と呼び、N をフレーム長、T をフレーム間隔と呼ぶ。

(2) 窓関数をかける:ある有限区間以外で 0 となる関数であり、フレーム化されたデータに対して重みをつける関数である。フレーム化処理を行う場合、離散的なデータの繋ぎ目においての信号の急激な変化の影響を和らげるため、原則として窓関数をかけなければならない。代表的なものとして音声信号だけに有効なハニング窓と、音声信号以外にも様々な信号にも有効なハミング窓がある。

$$\text{ハニング窓} : \omega(n) = 0.5 - 0.5\cos(\frac{2\pi n}{N-1}) \qquad (n = 0, 1, \cdots, N-1) \tag{2.2}$$

$$\text{ハミング窓} : \omega(n) = 0.54 - 0.54\cos(\frac{2\pi n}{N-1}) \qquad (n = 0, 1, \cdots, N-1) \tag{2.3}$$

(3) スペクトル分析（離散時間フーリエ変換、高速フーリエ変換):フレーム化処理によって得られた信号系列の短時間フーリエスペクトルは、離散時間フーリエ変換により以下の式で与えられる。

$$S(n) = \sum s_\omega(n)e^{-j2\pi\frac{nk}{N}} \qquad (k = 0, 1, \cdots, N-1) \tag{2.4}$$

離散フーリエ変換 (DFT) は、離散的なデータをフーリエ変換する際に、通常のフーリエ変換の無限区間積分を有限の和で書き換えたもので、時間領域、周波数領域ともに離散化されたフーリエ変換のことであり、時間領域の表現を周波数領域における表現に変換する。また、逆に周波数領域の表現を時間領域の表現に変換する、つまり元の音響データに戻す変換を離散フーリエ逆変換 (IDFT) と呼び以下の式で与えられる。

$$S(n) = \frac{1}{N}\sum S(k)e^{j2\pi\frac{nk}{N}} \qquad (k = 0, 1, \cdots, N-1) \tag{2.5}$$

実際の信号処理過程では、離散的フーリエ変換 (DFT) をその高速算法である高速フーリエ変換 (FFT) を用いて実行し、当該音声区間のスペクトル表現とすることが一般的である。高速フーリエ変換は式 (2.2)、(2.3) の N が 2^n 個であるとき、その処理を高速にできる性質がある。フーリエ変換の式には、

$$S\prime(n) = S(e^{j\frac{2\pi}{n}k}) = \sum s_\omega(n)e^{-j2\pi\frac{2\pi}{N}kn} \qquad (k = 0, 1, \cdots, N-1) \tag{2.6}$$

なる複素系列 $S\prime(k)$ が音声スペクトル表現として最も一般的に用いられる。

(4) パワースペクトルの算出:音響信号の特徴は主として、調音フィルタの振幅伝達特性に含まれる。したがって、音響信号の振幅スペクトル $|S'(k)|$、あるいはその 2 乗のパワースペクトル $|S'(k)|^2$、対数スペクトル $10\log|S'(k)|^2$ が注目すべきスペクトル表現になる。このスペクトルをグラフにすることで、音響データに含まれている周波数成分を解析することができる。音響データに含まれている周波数成分の上限はサンプリング定理により、サンプリング周波数の半分なので、高速フーリエ変換における周波数成分が意味のあるスペクトルとして扱われるのは 0 から π までのスペクトルである。また、音響信号の離散パワースペクトル系列は、離散スペクトル系列から式 (2.7) で表される。

$$|S'(k)|^2 = \frac{1}{N}[\mathrm{Re}\{S'(k)\}^2 + \mathrm{Im}\{S'(k)\}^2] \tag{2.7}$$

この 2 乗値のパワースペクトル $|S'(k)|^2$ を特徴量として扱っている。音響信号に高速フーリエ変換を施すと、時間表現 (縦軸:パワー、横軸:時間) から周波数表現 (縦軸:振幅、横軸:周波数) へと変換できる。しかし、実際には縦軸を周波数、横軸を時間としたグラフがよく使用されており、このようなグラフをスペクトルグラムという。スペクトルグラムは音声を視覚化したものであり、声紋とも呼ばれる。

2.1.2 音響特徴パラメータ

本研究で使用する 7 つの音響特徴パラメータについて述べる [10]

(1) スペクトルの変化

動的特徴量を連続するスペクトルのフレーム間の変化量として取り出す。音響信号のスペクトル分析した連続するフレームにおいて、あるフレームとその一定時間後のフレームとのパワースペクトルの差分によりスペクトルの変化量を得て、そのスペクトルの差分を一定時間足し合わせたものとしている。スペクトルの変化量によって比較する利点は、音声の識別に有利であり音声に比べて背景雑音のほうがスペクトルの変化量が大きく、無音のほうがスペクトルの変化量が小さいということである。

(2) スペクトルの傾き

あらかじめ人手により作成したラベルにより音響データの各区間を各種別（音声、音楽、背景雑音、無音）に振り分け、それぞれに対してスペクトル分析を行い、パワースペクトルを取り出し、各種別内において集められたパワースペクトルの分布を求めることで各種別において傾き値を得る。この傾き値を基に、与えられた音響ファイルから次々に得るパワースペクトルと各種別の学習データとの特徴パラメータの分布の類似度を比較する。この最小単純形は、パワースペクトルにおける一次回帰直線の傾きを比べることと同じである。傾きによって比較する利点は、有色系の音のほうが白色雑音よりも傾きが大きいので、音声と音楽と無音の識別に有利である。

(3) 白色雑音との近さ

パワースペクトルより一次回帰直線からスペクトル波形の切片を求めることで入力信号の白色性の度合を計測する。この白色雑音との近さによって比較する利点は、背景雑音のような定常的に混入した雑音は白色性が高いので、これらの識別ができることということである。

(4) ピッチ

有声音源の繰り返し周期、いわゆるピッチ（基本周波数）の変化を調べることで、音源の変化を知ることができ、音源の特定のパラメータである。周波数分析によりピッチを求め、学習データと比べることで音源の特定に用いる。

(5) パワー

時間領域の分析だが、音響信号のような非定常的な信号に対して、変化していく信号の大きさにうまく追随するような比較的短い区間に音響データを区切り、その区間の信号 $xl(n)$ に対してエネルギー $E(l)$ を定義する [11]。

$$E(l) = \sum_{n=0}^{N-1} \{x_l(n)\}^2 \tag{2.8}$$

ここでは、整数 N は窓の中に含まれる音響信号の数である。

利点としては、測定が簡単であり、音声認識における有色系の音の区間の抽出にもよく用いられることから、有音と無音の区別に有利である。

(6) 中心周波数

抽出したパワースペクトルにおいて、無音の場合は右下がりに傾斜しているが、有音の場合は傾斜の途中で膨らみまたは突起が発生する。その突起がもっとも大きく発生している周波数帯の中心部分の周波数を中心周波数として定義している。これは有音と無音の識別に効果がある。

(7) 中心周波数のバンド幅

中心周波数を含む膨らみ、あるいは突起の始まりと終わりによる周波数帯の長さをバンド幅として定義する。音声は一定の周波数を含むことが多いためそのバンド幅はある程度の大きさになることが考えられるが、雑音はあまり多くの周波数を含まないものから白色性が高く幅広い周波数を含むものまで様々であり、その違いから音声と雑音の特定に有効である。

2.2　i-vector の概要

近年の話者認識システムの多くは i-vector [3][4] に基づいて構成されおり、この領域における最高水準の技術となっている。i-vector とは、ある発話から得られた音響特徴量を因子分析を用いて、話者固有の特徴を抽出したものである。i-vector の抽出においては、因子分析の入力として、発話毎に GMM(Gaussian Mixture Model) の平均ベクトルを結合した GMM スーパーベクトルを用いる。発話 u から作成された GMM スーパーベクトル $M_u \in R^{CD_F}$ は以下で定義される。

$$M_u = Tw_u + m \tag{2.9}$$

ここで M_u は大量の不特定話者の発話データから作成される UBM（Universal Background Model）を事前情報として事後確率最大化（MAP）法により推定された GMM を用いる。また m は UBM から得られる話者及びチャネル非依存の GMM スーパーベクトルである。C は GMM (UBM) の混合数，D_F は音響パラメータの次元数、$T \in R^{CD_F * D_r}$ は低ランクの矩形行列 $D_r \ll CD_F$ で、全変動空間を張る基底ベクトルで構成される固有声行列である。$W_u \in R^{D_r}$ は発話ごとに与えら

れる潜在変数であり、平均ベクトルが $0 \in R^{D_T}$ で共分散行列行列が単位行列 $I \in R^{D_T * D_T}$ のガウス分布 $N(w; 0, I)$ に従う。この w は total factor(全因子) と呼ばれ、各発話に対する i-vector である。つまり、i-vector は GMM スーパーベクトル空間における平均的な話者 (UBM の平均) から「差 (を次元圧縮したもの)」として各話者を表現したものと言える。

2.2.1 UBM に対する Baum-Welch 統計量

準備として、UBM に対する Baum-Welch 統計量を計算することから始める。$O_u = o_1, o_2, o_3, \cdots, o_L$、$o_t \in R^{D_F}$、を発話 u から得られる L フレームの音響パラメータ系列 $c = 1, 2, 3 \cdots, C$、を UBM (GMM) の混合要素を表す添え字、$\Omega = \{\pi_c, m_c, \sum_c\}_{c=1}^{C}$ を UBM のパラメータ (混合重み、平均ベクトル、対角共分散行列) とする。このとき、発話 u に対する 0 次、1 次、2 次の Baum-Welch 統計量は、

$$N_{u,c} = \sum_{t=1}^{L} \gamma_t(t) \tag{2.10}$$

$$F_{u,c} = \sum_{t=1}^{L} \gamma_t(c)(o_t - m_c) \tag{2.11}$$

$$S_{u,c} = diag\left[\sum_{t=1}^{L} \gamma_c(c)(o_c - m_c)(o_t - m_c)'\right] \tag{2.12}$$

と書ける。ここで、$y_c(c)$ は、o_t が UBM の c 番目の要素分布から生成される事後確率

$$\gamma_c(c) = p(c \mid o_t, \Omega) = \frac{\pi_c p(o_t \mid m_c, \sum c)}{\sum_{k=1}^{C} \pi_k p(o_t \mid m_k, \sum k)} \tag{2.13}$$

である。更にこれらを用いて、

$$N_u = \begin{pmatrix} N_{u,1}, I_{D_F} & 0 & \dots & 0 \\ 0 & 0 & \dots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & N_{u,C}, I_{D_F} \end{pmatrix} \in R^{CD_F \times CD_F} \tag{2.14}$$

$$N_u = \begin{pmatrix} F_{u,1} \\ F_{u,2} \\ \vdots \\ F_{u,C} \end{pmatrix} \in R^{CD_F} \tag{2.15}$$

$$N_u = \begin{pmatrix} S_{u,1}, 0 & 0 & \dots & 0 \\ 0 & S_{u,2} & \dots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & S_{u,C} \end{pmatrix} \in R^{CD_F \times CD_F} \tag{2.16}$$

ここで、$I_{D_F} \in R^{D_F \times D_F}$ である。

2.2.2 全因子 w の確率分布と i-vector の抽出

本節では、w に関する種々の確率分布を導出する。このとき、w の事後分布の導出過程において i-vector の具体的な計算方法を示す。

- 事前分布

  w の事前分布は $p(w)$ 平均 0、共分散行列を持つガウス分布であり、以下のように書ける。

$$p(w) = \propto exp(-\frac{1}{2}w'w) \tag{2.17}$$

- 条件付き分布

  $M_{u,c}$ を混合要素に対する M_u の部分ベクトルとする。直感的には、$M_{u,c}$ は発話 O_u で学習した GMM の混合要素 c に割り当てられた O_c の各フレームは、平均 $M_{u,c}$、共分散行列 $\sum_c$(UBM のまま) に従うと仮定する。すなわち、w の値で条件付けられた観測データ O の条件付き分布は

$$P(O_u|w_u) = etp\left(\sum_{t=1}^{L}\sum_{c=1}^{C}\gamma_t(c)log(2\pi)^{-\frac{D_F}{2}}\left|\sum_{c}\right|^{-\frac{1}{2}} - \frac{1}{2}\sum_{t=1}^{L}\sum_{c=1}^{C}\gamma_t(c)D(o_t;\theta_c)\right) \tag{2.18}$$

のように書ける。ここで、

$$D(o_t;\theta_t) = (o_t - M_{u,c})'\Sigma_c^{-1}(o_t - M_{u,c}) \tag{2.19}$$

$$M_{u,c} = m_c + T_c w_u \tag{2.20}$$

である。$T_c \in R^{D_F \times D_T}$ は、混合要素 c に対する T の部分行列である。式 (2.18) の exp の内部を Baum-Welch 統計量を用いて整理すると、

$$\sum_{t=1}^{L}\sum_{c=1}^{C}\gamma_t(c)log(2\pi)^{-\frac{D_F}{2}}\left|\Sigma_c\right|^{-\frac{1}{2}} - \frac{1}{2}\sum_{t=1}^{L}\sum_{c=1}^{C}\gamma_t(c)D(o_t;\theta_c) = G_u^{\Sigma} + H_u^{\Sigma T} + Const \tag{2.21}$$

ここで、G_u^{Σ} 及び $H_u^{\Sigma T}$ は、

$$G_u^{\Sigma} = \sum_{c=1}^{c}\left[\frac{1}{2}N_{u,c}log\left|\Sigma_c^{-1}\right| - \frac{1}{2}tr\left(\Sigma_c^{-1}S_{u,c}\right)\right] \tag{2.22}$$

$$H_u^{\Sigma T} = w_u'T'\sum^{-1}F_u - \frac{1}{2}w_u'T'N_u\sum^{-1}Tw_u \tag{2.23}$$

- 事後分布

  w の事後分布は (2.18)〜(2.23) を用いると、

$$
\begin{aligned}
p(w_u|O_u) &\propto p(O_u|w_u)p(w_u) \propto exp(w_u'T'\sum^{'} Tw_u - \frac{1}{2}w_u'w_u) \\
&\propto exp(w_u'T'\sum^{'} F_u - \frac{1}{2}w_u'N_u\sum^{-1} Tw_u - \frac{1}{2}w_u'w_u) \\
&\propto exp(-\frac{1}{2}(w_u - G_uT'\sum^{-1} F_u)'G_u^{-1}(w_u - G_uT'\sum^{'} F_u))
\end{aligned}
\tag{2.24}
$$

と書ける。ここで、

$$
G_u = (I + T'\sum^{-1} N_uT)^{-1} \tag{2.25}
$$

である。w の事後分布もガウス分布であることに注意すると、平均及び分散は、

$$
E[w_u] = G_uT'\sum^{-1} F_u \tag{2.26}
$$

$$
cov[w_u] = G_u \tag{2.27}
$$

となる。前述のとおり、確率的潜在変数モデルのもと、i-vector は w の事後分布の平均として得られる。つまり、発話 u の i-vector は、Baum-Welch 統計量 N_u、F_u 及び推定済みのパラメータ T,Σ を用いて、式 (2.26) により計算することができる。

2.2.3 因子分析モデルパラメータの推定

因子分析モデルのパラメータ T 及び Σ は、EM アルゴリズムにより求められる。すなわち、完全データ $(O_u, w_u)_{u=1}^{U}$ に対する対数尤度の期待値

$$
Q = \sum_{u=1}^{U} E[logp(O_ww_u|\theta)] \tag{2.28}
$$

の最大化問題を解くことで求める。ここで、θ はパラメータ T、Σ を表す。完全データの対数尤度は、

$$
logp(O_ww_u) = logp(O_u|w_u, \theta) + logp(w_u) \tag{2.29}
$$

と書けるので、式 (2.18)〜(2.23) を用いると、式 (2.28) は以下のように整理できる。

$$
\begin{aligned}
Q = \frac{1}{2}\sum_{u=1}^{U}\sum_{c=1}^{C}(N_{u,c}log\left|\Sigma_c^{-1}\right| - tr(\Sigma_c^{-1}S_{u,c})) \\
+\sum_{u=1}^{U} tr\left(\Sigma^{-1}\left(F_uE[w_u']T' - \frac{1}{2}N_uTE[w_uw_u']T'\right)\right) \\
-\sum_{u=1}^{U}\frac{1}{2}tr(E[w_uW_u'])
\end{aligned}
\tag{2.30}
$$

以上より、E ステップにおいては古いパラメータを使って、w 空間の事後分布の統計量を以下のように計算する。

$$E[w_u] = G_u T' \Sigma^{-1} F_u \tag{2.31}$$

$$E[w_u w'_u] = G_u + E[w_u]E[w'_u] \tag{2.32}$$

M ステップでは、式 (2.30) をパラメータに関して最大化する。まず、(2.30) を T に関して微分して 0 と置くことで、以下の関係式を得る。

$$\sum_{u=1}^{U} \Sigma^{-1} F_u E[w'_u] = \sum_{u=1}^{U} \Sigma^{-1} N_u T E[w_u w'_u] \tag{2.33}$$

これより、T の推定式が、

$$T^i = \left(\sum_{u=1}^{U} F_u^i E[w'_u] \right) \left(\sum_{u=1}^{U} N_{u,c} E[w_u w'_u] \right)^{-1} \tag{2.34}$$

のように得られる。ここで、T^i、F_u^i は、おのおの T、F_u の i 行目を表し、$i = (c-1) \times D_F + f$、$1 \leq f \leq D_F$ である。また、Σ の推定式は、

$$\Sigma = N^{-1} \left(\sum_{u=1}^{U} S_u - diag \left[\sum_{u=1}^{U} F_u E[w'_u] T' \right] \right) \tag{2.35}$$

となる。ここで、$N = \Sigma_{u=1}^{U} N_u$ である。

2.2.4 コサイン類似度

発話 x から抽出した i-vectorw_x と発話 y から抽出した i-vectorw_y の比較を行うための方法としてコサイン類似度を用いる。

$$\cos(w_x, w_y) = \frac{w_x \cdot w_y}{\| w_x \| \| w_y \|} \tag{2.36}$$

類似度の値の範囲は、$-1 \leq \cos(w_x, w_y) \leq 1$ であり、類似度が最も高い値は 1 である。

2.3 i-vector を用いたアンカーの発話区間抽出手法

クラスタリングについて

2.4　音声認識

2.4.1　音響モデル

2.4.2　言語モデル

2.4.3　単語辞書

2.4.4　DNN の概要

第3章

提案手法

ていあんしゅほう

図 3.1: ぴあの

図 3.2: ぴあの

第4章

評価実験

ひょうかじっけん

4.1 予備実験

4.1.1 実験目的

しきいちきめるよー

4.1.2 実験条件

おおざっぱだよー

4.2 本実験

4.2.1 実験目的

ひょうかするよー

4.2.2 実験条件

よびといっしょだよー

表 4.1: 実験結果

被験者	実験 1	実験 2	実験 3
AAA	3	5	3
BBB	1	2	4
CCC	5	5	6
DDD	5	0	8

4.3　考察

こうさつするよー

第5章

結論

けつろん

第6章

追加

追加

謝辞

最後に，本研究および本修士論文作成にあたり暖かい御指導および適切な御助言をして頂いた松永　昭一教授，また，関係者各位に心より感謝いたします．

また，同研究室博士前期 (修士) 課程 2 年の博士前期 (修士) 課程 1 年の学士課程 4 年のその他関係各位に心から感謝いたします．

参考文献

[1] Google