

エイブルサポート系 知的画像処理システムチーム

陳 延偉

(情報理工学部教授)

○陳 きょうはどうもありがとうございます。情報理工学部の陳と申します、よろしく願いいたします。

私の方は理工系ですので、恐らく皆さんと若干専門分野が違いますので、ぜひ皆さんの方からまた御不明な点とか御批判などいただければと思っています。

まずはプロジェクトについて、キーワードは対人援助ということですので、じゃあ我々理工系から見た対人援助、私個人ととってもいいのですけれども、対人援助とはどういうものなのかについて、少し初めにまずお話をさせていただきます。

理工系というのは、御存じのように基本的には物づくりがメインですので、物づくりという観点から見れば、対人援助というのはどういうものかということについても少し話をさせていただきます。

1985年にノースカロライナ州立大学のロナルド・メイス博士がユニバーサルデザインという概念を提唱し、注目されています。ユニバーサルデザインとは、どんな人でもたとえ障害者でも利用することのできる製品とか建物とか環境をデザインすることをいいます。恐らくユニバーサルデザインのベースになっているのがパブリック・ユーズですね。要はみんなが使うものに限っては、そういうユニバーサルデザインが必要ですが、一方例えば個人用ですね。パーソナル・ユーズに限ってはどうかというと、むしろその個人、ユーザーに合ったものがデザインできれば最もいいんじゃないかということで、そこで私が考えたのはパーソナル・オリエンテッド・デザインというものです。ここではちょっとそういう言葉を使ってるのですけれども、

これはどういうことかといいますと、個人の嗜好とか趣味とか、または習慣などに合った機器の設計とか、また手法の開発などを行うことです。

そこで我々のプロジェクトというか、我々のチームの目的ですね、どういう目的かということは、先ほど申し上げたように、今まで例えば私たち自身の専門は画像処理ですので、画像処理用としてPhotoshopというそうソフトが非常に有名ですが、Photoshopで画像処理すると、ユーザー、例えばAさんにしてもBさんにしても、基本的には大体同じような結果が得られます。方法が決まっていますので、それは同じような結果が出ていますが、しかし一方、例えばピカソの絵とか、好きな人もいれば、ちょっとわかりにくいと思う人も当然いますので、そういった芸術作品、または自分の感性を生かした画像処理を考える場合、やはりユーザーというか、処理する人の主観的な意識、または感性を考慮した、そういったような画像処理、またはそういった機器やシステムを開発する必要がある。そういったような人間の主観的な意識、または感性を考慮した知的画像処理システムを開発しようというのが一つの目的でございます。一応我々のプロジェクトのメンバーとしては、一応私がプロジェクトリーダーですが、共同研究者として蘇州大学の心理学専門家でもある黄先生と一緒に共同研究させていただいています。あとは私どもの学生や院生も一緒にやっております。

研究経過、私は昨年度からこのプロジェクトに参加させていただいて、大体今までどういうことを行ってきたかについて、ここに簡単に書いてあります。今、先ほど申し上げたこの研究の目的は、そういった人間の中の主観的な意識、または感性を取り入れた画像処理を中心に研究を行ってきました。例えば昨年度人間の主観を取り入れられる、インタラクティブ魚群システムを作成したり、または画像の特徴、主に色ですけども、色と人間の印象空間の写像関係どういうふうになっているか、特に年代別ですね、老人、成人、それから学生と子供ですね、そういう年代別で写像関係を求めました。そういったものを意識して研究を行ってきました。

それから今年度、19年度においてはカラー画像の特徴の人間の印象または心理空間への写像を我々が実際ニューラルネットワーク使って求めましたの

で、その応用としてそういった生地ですね、例えば私が明るい生地を服にしたいと、じゃあどういうものがあるかですね、そういったものが、自分の印象をキーワードとして入力するだけで、そういったような生地が検索できるシステムを、作成いたしました。後でちょっと簡単な例をお見せいたします。

そしてもう一つは今やっているのは、これから株式会社花王との共同研究なんですけれども、多視点または多照明の顔、アピアランス、見えですね。どのように変わっているか、とりあえず今年度後半からスタートしたばかりですが、そのアピアランスの統計的なモデル解析をしております。将来的にはこれら人間の印象に対してどういうふうに結びついていくかなどをやっていこうと思っています。だから来年度はそういうパーソナルマッピングファンクション、すなわち写像関数ですね、応用した画像検索システムをつくってみようというふうに今考えております。そういった顔のアピアランスと印象とのそういった写像関係とか解析もやっていこうと考えております。

じゃあまずきょうは主に二つの内容について、簡単に御報告をさせていただきます。まず画像の特徴と人間印象との写像関係及び生地検索システムの応用という内容について報告します。まず背景として画像の特徴ですね。我々は画像屋さんですので、まず画像の物理的特徴はどういうものなのか。例えば色とか、テクスチャーですね、テクスチャーというのは模様ですね、そして見え方というか形ですね。そういったものがどういうふうに数学的な手法を使って定量化して抽出すると。そして、あと心理空間の特徴ですね。これは印象と考えていただければいいかと思いますけれども、それに対してどのような印象を持つのか、心理空間にどのように写像されているかというもの。そして実際にニューラルネットワークを使って、その写像関数（マッピングファンクション）を求めました。さらにそういったマッピングファンクションを使ってそういった生地の検索システムを作成しました。

これまで画像検索というのは大体キーワードを使って、キーワードを入力して画像を検索するのが一番本当にポピュラーなんですけれども、90年代後半からキーワードじゃなくて画像のコンテンツですね、画像のコンテンツは色とか形状とか模様など、コンテンツを使って画像を検索するというものでござ

います。一方、人間が好み、例えば私が明るい画像を検索したいとか、そういうふうな意向も当然ございますので、そういった人間の感性、またはそういった嗜好に合った画像を検索するというものが今は研究の対象になっています。その人間の感性はさまざまな条件によって異なりますので、そういったさまざまな条件に合った人間の感性を用いた画像技術システムの実現は、我々は目指しております。

そしてそれの大体のイメージでございませうけれども、ある画像を、画像については先ほど申し上げましたように、いろんな色がありますし、いろんな模様があって、またはいろんな形状がございませう。そういったような、これはどちらかというと物理的な空間でございませう。一方、画像を人が見ると、様々な印象をもちます。それを印象空間または心理空間への写像となります。子供の印象と老人の印象がまたことなるので、同じ画像特徴空間上の点ですが、心理空間に写像すると、年代などの条件によって、異なる点に写像されます。そのような年代別の写像関数（マッピングファンクション）は昨年度の研究において、求めました。今年度はそういう写像関数（マッピングファンクション）を用いて実際に心理空間における人間の印象とか感性で画像検索するというシステムを作成いたしました。ちょっと余り時間がございませうので簡単に説明いたします。

一応画像の特徴ということですね。例えば様々な生地画像をサンプルとして集めていきまして、それでそれに対してそれぞれの生地の色情報はどのようなものなのか、模様情報はどのようなふうなものなのか、細かいのか、粗いのか、それぞれ特徴を定量化して、定量的に抽出して評価します。それで逆にまたそれぞれの画像に対して、サンプル画像に対して、12名の大学生にアンケート調査を行いました。これはSD法を用いてサンプル生地画像に対する23項目の印象を5段階評価で行いました。23個の印象項目はここに示していますが、不思議な・不思議でない、明るい・暗い、おもしろい・退屈、そういったような23個の印象項目（形容詞ペア）を選んで評価してもらいました。そして色の特徴としてRGBの値を用いることができますが、本研究ではRGBをまずHSV空間に変換し、Hという値のみを用いて色特徴量としまし

た。例えばこういった色をHSV空間に変換し、Hの値を表示しますと、こういう感じになります。一方こちらの画像は割と緑色に集中していますので、ほぼこの辺に値が集中して出ているという形です。そして模様の方は、一定の間隔で、すなわち一定の周波数の模様でございますので、それもフーリエ変換を利用して、周波数とその方向を定量的に求め、テクスチャの特徴量としました。次に見え方すなわち形状について紹介します。我々は様々な方向性をもつGaborフィルタを用いて、いろんな方向のエッジを検出し、形状の定量評価を行いました。たとえば、この画像に対して、横方向のGaborフィルタを用いると、横のエッジ画像のみが得られます。上述のような方法で色、テクスチャ（模様）及び形状の特徴量を抽出し、それぞれの特徴量に対して主成分分析（PCA）を用いて次元圧縮を行いました。画像の特徴量はマッピングファンクションを学習するための入力として用いる。

上述のような色情報、または模様情報、または形状情報に対して、人間がどのような印象を持つかの対応関係をニューラルネットワークを用いて近似する。画像特徴量（色、テクスチャ、形状）はニューラルネットワークの入力として用いる。ネットワークの出力は23項目の印象になります。ニューラルネットワークの学習には、SD法で得られた23項目の印象を教師信号として用いました。学習したニューラルネットワークを用いて、テスト画像の印象を予測した結果はここに示します。比較のため実際SD法で得られた印象も同図に示します。両方を比較すると、非常によく予測されていることがわかります。すなわち、画像特徴空間から人間の心理空間への写像関数（マッピングファンクション）がよく求められたことになります。

求められた写像関数を用いて、図のような感性に基づく画像検索システムを作成しました。検索したい画像の印象を入力するだけで、目的の画像が出力されます。例えば明るいところで検索すれば明るい生地が出てきます。そして、暗いところでやれば、こういった暗い生地が出てきます。そして例えば暗くて細かい条件を入力すると、こういうものが出てきます。それに対して、検索した結果に対して実際に学生に見てもらい、目標に合っているかどうか評価してもらいました。例えば明るい画像の検索結果に対して、やはり

明るい画像だと感じる人が非常に多いです。これが濃い画像に対する評価結果ですね。やはり本システムが検索した結果と人間が一致していることがわかります。

ちょっと最後に簡単に顔のアピランス評価に関してちょっと話をします。時間があまりありませんが。

これは株式会社花王との共同研究でもございますけども、人間の顔は化粧の仕方や照明の当て方、または視点によっては見え方が大分違いますよね。我々はそのような顔の見え方を統計的に解析し、このような顔の質感関数を出そうというのが目的です。まず我々と花王の共同で、KAO-Ritsumeikan Multi-angle View, Illumination and Cosmetic Facial Image Database（略してMaVIC）という多視点多照明化粧顔データベースを構築しています。構築したデータベースはMaVIC 1、MaVIC 2、MaVIC 3 という三つから構成されています。ここに示すようなデータが入っております。画像は株式会社花王の方で作成した装置で撮影されたものです。15度間隔でさまざまな視点から撮影できて、そして照明も同じように15度間隔で照らすことができます。そしてここにプロの美容師で化粧してもらったシャープな顔、ソフト顔などです。我々はまずPCAを用いて視点ごと、照明ごとの固有空間を計算し、さまざまな化粧顔をその固有空間に投影します。固有空間での座標、すなわち統計モデルの係数をプロットしたのはこの図であります。図に示すように、化粧仕方によって、各成分の係数が大きく変わっています。特に提示モードにおいて、すなわち、統計モデルの係数を用いていると、顔の化粧仕方（質感）を定量的に表すことができます。それぞれの成分の物理意味は現在検討中です。今後は顔の質感と人間の印象との関係についても調べていく予定です。

以上簡単でございますけども、我々の研究をここで発表させていただきました。ありがとうございます。