

# 大量のDSC撮影画像からの自動選択を可能とする 画像評価技術“iAgent”の開発

野中 俊一郎\*, 松井 優子\*, 内田 充洋\*, 羽田 典久\*

## “iAgent”, an Image Recommendation/Rating Technique Enabling Automatic Selection among a Large Number of DSC Image

Shunichiro NONAKA\*, Yuko MATSUI\*, Mitsuhiro UCHIDA\*, and Norihisa HANEDA\*

### Abstract

Most of captured DSC image data are not outputted with only few being printed in the form of hardcopy. One reason is the time and labor required for selecting frames to be printed. Thus, to promote the use of digital images, we have developed a new image rating technique called “iAgent”. First, DSC users were classified. Among various classes, a class for young family having the youngest child before school age and being the most active DSC user class was chosen in the present study. Their behavior of selecting frames to be printed was analyzed by direct observation and interviews. Basic four steps were extracted and logically implemented into “iAgent”. More than 60% of users are satisfied with the selection result of “iAgent”.

### 1. はじめに

近年、デジタルカメラ（以下、DSC）の普及に伴い、一般ユーザが撮影する画像のショット数が増大している。一方、撮影された画像の内訳を見ると、従来よりもメモ的な撮影や同一シーンでの念のための撮影などが増えており、総撮影ショット数におけるユーザがプリントなどの利活用を望むショット数が占める比率は大きく低下している。また、DSCを使用した場合、DSC付属の液晶画面もしくはPCの画面などで画像の内容を確認できるため、相対的にプリントの必要性が低下している。

われわれは、このような状況に鑑み、デジタル画像の利活用を促進するための技術として“iAgent”を開発した。iAgentは、デジタル画像の「イベント」「類似性」「人物の顔の情報」「明るさ」「ボケ・ブレ」「色」の観点から、ユーザにとってのデジタル画像の価値を自動評価する技術である。iAgentは、氾濫するデジタル画像の中からユーザが利活用に適した画像を自動的に選択できる画期的技術であり、デジタル時代の画像の利活用における必須技術として位置づけられる。

本報告では、iAgentの開発の背景、開発の手順、技術の詳細と効果、今後の課題について解説する。

### 2. iAgent開発の背景

近年、DSCの普及に伴い、カメラ1台あたりの撮影ショット数は銀塩カメラ時代と比較して約3倍に増大している。しかし、一方で撮影した画像をただ保存しておくのみで積極的に利活用しないいわゆる「死蔵」されるショット数も増大しており、DSC時代にふさわしい画像の利活用が十分なされているとはいいがたい状況である。

この「死蔵」の大きな原因となっているのが、利活用に好適な画像を選択する手間である。膨大な画像から目的に合致した画像を平易に選び出す手段が事実上存在しないため、画像の利活用が阻害されているのである。DSC撮影画像の利活用に関するユーザ行動の例として、DSC撮影画像のプリント行動に関するユーザ調査結果をFig. 1に示す。Fig. 1からわかるように、プリント意向があるユーザは全体の約7割を占めるが、そのうち約半数が「プリントすべき画像を選択する手間」の障壁のためにプリントを行なわないでいる。

この状況を打破するために、ユーザに立ち代った画像の自動選択という機能提供が必要である、との仮説を立て、われわれはiAgentの開発に着手した。

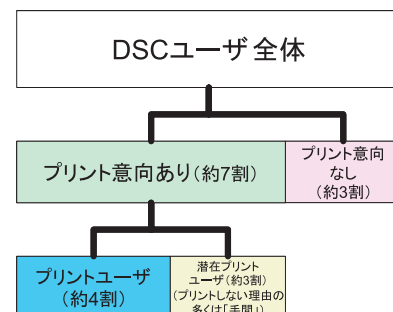


Fig. 1 Classification of DSC users.

本誌投稿論文（受理2006年12月4日）

\*富士フイルム（株）R&D統括本部  
ソフトウェア開発センターインキュベーション部  
〒351-8585 埼玉県朝霞市泉水3-13-45

\*Business Incubation Division  
Software Development Division  
Research & Development Management Headquarters  
FUJIFILM Corporation  
Senzui, Asaka, Saitama 351-8585, Japan

### 3. iAgent 開発の進め方

われわれは、画像の自動評価技術 iAgent の開発を決意したが、撮影者に立ち代って自動で画像に評価点を付けることは不可能な課題であると考えられていた。ユーザによって画像選択の基準は当然異なり、また、利用目的やほかの画像との相互関係によっても選択結果は異なるはずである。

そこで、iAgent を開発するにあたり、一般ユーザがプリント画像を選択するプロセスを観察し、その選択ロジックを自動化するという手順を取るとともに、対象とするターゲット層を絞り込んだ技術検討を行ない、めどがついた後に対象を広げていくという手法を取ることにした。今回の開発では、DSC の利用度が最も大きいヤングファミリー層（末子が未就学）をターゲットにした。

まず、ユーザがプリント画像を選択するプロセスについて行動観察と聞き取り調査を行なった。その結果、プリント画像選択のプロセスには共通の手順があることがわかった。その手順は、以下の通りである。

- 1) 画像を何らかの基準でグループ分けする
- 2) グループの中から重要なグループを選択する
- 3) 重要なグループに含まれる画像の中から失敗写真や不要な画像を候補からはずす
- 4) 候補に残った写真の中から好適な写真を選ぶ

また、上記手順の 1) のグループ分けは、まれに、「被写体」が基準として用いられる場合を除いて、ほとんどの場合、「イベント」を基準に階層的に分類されることがわかった。

次に、さらに詳細なプリント画像選択の基準について聞き取り調査を行なった。その結果、不適な画像を取り除いたり、好適な画像を選択したりする基準は、「重要な画像でない（メモ的な撮影である）」「特定の人物が写っている」「被写体の表情がよい」「被写体の人数」「顔の向き」「明るさ」「ボケ・ブレ」であることがわかった。

そこで、iAgent の開発においては、上記プロセスを自動的に再現することを目指し、DSC 撮影画像のグループ分けとグループの重要度の算出、およびグループ内の画像の適否の自動評価を行なうロジックの開発を目標とした。

約 100 名の対象ユーザから、約 3 万枚の DSC 撮影画像と撮影画像に対する評価（実際にプリントしたか、など）をヒアリングした結果を DB 化し、この DB の「正解評価結果」に最も近似した評価値を出力するよう機械学習などを用いて自動評価ロジックを構築する、というプロセスで開発を行なった。

## 4. iAgent の構成

### 4.1 全体構成

Fig. 2 に iAgent の構成を示す。

iAgent はソフトウェアライブラリとして用意され、PC 上での画像ビューワやプリント受付機上のアプリケーションなどから呼び出されて、複数の画像の入力に対し、それぞれの画像の評価値を出力する、という機能を持つ。

3 章で目標とした手順に則して、画像のグループ分けのために、「(撮影時刻による) イベント分類」と「類似画像分類」のモジュールを用意し、グループ内の画像の

適否を評価するために、「イベント重要度評価」「顔評価」「明るさ評価」「ボケ・ブレ評価」「色評価」のモジュールを用意した。さらに、「イベント重要度評価」「顔評価」などの個々の観点からの評価値をユーザの好みなども踏まえて総合的に評価するための「総合評価」モジュールと、全体の処理の制御、および画像ファイルの読み込み・処理結果の入出力を行なう「iAgent 制御」モジュールを設ける。

以下、個々のモジュールの詳細について説明する。

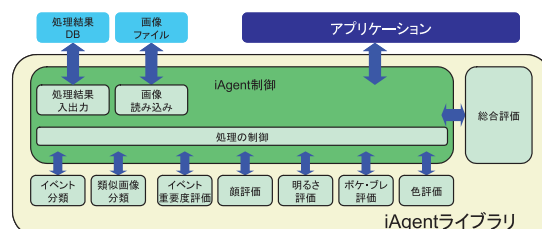


Fig. 2 Structure of "iAgent".

### 4.2 イベント分類・重要度評価

3 章でも触れたように、プリント画像選択のプロセスの最初に、「画像を何らかの基準でグループ分けする」という行動があることがわかっている。この「グループ分け」は、ほとんどの場合、「イベント」を基準に行なわれている。ユーザ撮影の DSC 画像を解析したところ、撮影画像は「イベント」という概念で確かに分類でき、撮影時刻をもとに分類できそうであることがわかった。そこで、われわれは DSC 画像に記録されている複数の画像の撮影時刻をもとに「イベント」の範囲を推定し、画像を階層構造のイベントに自動分類する技術を開発した。

われわれは、「同一イベント内の“撮影”という事象は、ポワソン分布に従い発生する」との仮説を立て、撮影間隔がポワソン分布の予想する範囲を超えた場合にイベントの分割を行なう、という考え方でイベント分類を行なうソフトウェアを開発した。また、分類されたそれぞれのイベントについて、階層の深さ・画像数・類似画像数・時間・人物数をもとにイベント重要度を算出した。

その結果、Fig. 3 に示すように、撮影者の意図に合ったイベント分類が実現できていることがわかった、また、算出したイベント重要度は、実プリント数との相関係数が 0.83 となり、有用であることが確認できた。

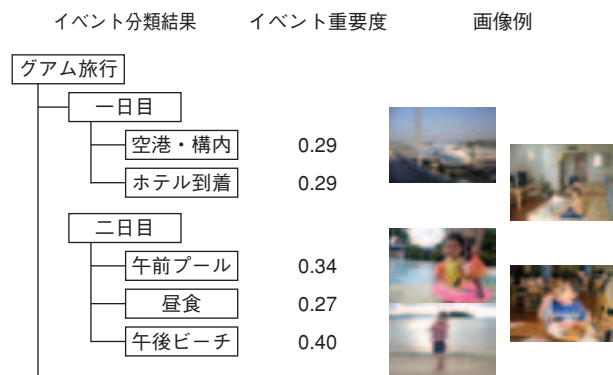


Fig. 3 Example of event categorization.

### 4.3 類似画像分類

DSCの時代となり、ユーザの撮影行動が大きく変化したことの一つに、「類似画像の撮影がきわめて増加した」ことが挙げられる。DSCユーザ調査の結果、Fig. 4に示すように、DSCユーザ撮影画像には平均38%の類似画像が存在し、25%はほかに類似画像がある重要度が低い画像であることがわかった。

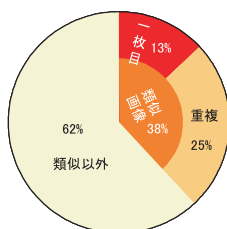


Fig. 4 Occurrence of similar frames in DSC-capturing data.

通常の場合、類似画像のグループからは1枚の画像が選択されれば十分であると考えられることから、DSC撮影画像同士の類似性を自動判定し、イベント分類後の画像をさらに類似画像の観点で分類する技術を開発することにした。

まず、任意の2枚の画像が類似であるか否かを判定するロジックを開発することにし、その開発のための学習データとして、ユーザから集めた約3万枚の画像から約1万組ずつの類似画像ペアと非類似画像ペアを用意した。画像ごとに、位置に応じた色・明るさやそれらのヒストグラムなどを特徴量として求め、これらの特徴量をもとに学習データを使ってAdaBoost<sup>1)</sup>でチューニングした判別機を構成し、類似画像判定ロジックとした。

その結果、ユーザ撮影画像の類似画像ペアの類似判定成功率が約93%という、今回の目的には十分な精度を達成した。Fig. 5に示す類似画像判定の成功例からも、回転やズーム、画角の違いなど、日常撮影に起こりうるさまざまな揺らぎに対応した実用的な類似判定が実現できていることがわかる。このモジュールは、ほかの用途にも十分応用可能である。



Fig. 5 Example of frame-to-frame similarity.

### 4.4 顔評価

今回、われわれがターゲットとしたヤングファミリー層は、子供や家族の記録が主要な撮影のテーマであり、「顔」が画像評価の重要な要素であることが容易に想像できる。そこで、本研究では既開発の顔検出モジュール

ル<sup>2)</sup>から得られる顔情報を用い、顔の観点からの画像評価を行なった。

具体的には、ユーザ画像の中から評価が高い画像を解析し、画像中の人物数・顔の位置・顔の大きさ・顔の向きをもとに、良好な画像となりうる組み合わせを定式化した。例えば、Fig. 6のように、顔が中央あるいは黄金分割の位置にある、などの画像の評価が高くなり、複数の顔の向きがバラバラ、などの画像の評価が低くなる。その結果、相関係数0.91とかなり満足のいく結果を得た。

| 被写体 2 人                          | 被写体 3 人                       |
|----------------------------------|-------------------------------|
| <p>0.49<br/>※顔の位置が黄金比</p>        | <p>0.50<br/>※顔の位置が中央+黄金比</p>  |
| <p>0.41<br/>※位置が悪く、また外を向いている</p> | <p>0.45<br/>※位置が端により過ぎている</p> |

※数値は顔評価値

Fig. 6 Evaluation of face-containing frames.

### 4.5 明るさ評価

プリント画像選択のプロセスのうち、特に、「不適な画像を除去する」際に重要な項目として「明るさ評価」がある。夜景などを含む画像一般において、「明るさ」から画像の評価を行なうことは困難であるため、本研究では4.4の顔検出モジュールにより検出された画像中の最大の顔領域の明るさを評価した。

その結果、Fig. 7に示すように、被写体の明るさを適正に評価し、アンダーや逆光などの不適画像を適切に除去することが可能となった。



Fig. 7 Frame evaluation in terms of brightness.



#### 4.6 ボケ・ブレ評価

4.5節の「明るさ評価」と同様に、特に、不適画像を除去するという観点では「ボケ・ブレ評価」が重要である。富士フィルムでは、ボケ画像を自動修正する技術を開発している<sup>3)</sup>。本研究ではこの技術を応用し、4.4節の顔検出モジュールにより検出された画像中の最大の顔領域のボケ・ブレを評価した。

その結果、Fig. 8に示すように、ボケ画像を適正に評価し、明らかな失敗写真を除去することが可能となった。



Fig. 8 Frame evaluation in terms of the degree of blur.

#### 4.7 色評価

4.4節、4.5節、4.6節にて「不適な画像を除去し、好適な画像を選択する」ためのモジュールを解説したが、これらはいずれも被写体が人物である場合に効果を発揮するものであった。本研究で対象とするヤングファミリー層のDSC撮影画像は、6割程度が人物を被写体とするものであるが、残りの4割程度の撮影画像は、風景などの人物以外を被写体とするものである。これらの画像の評価を目的に、画像の色を用いて画像を評価する「色評価」モジュールを開発した。

画像中の局所的な彩度・輝度の分布とヒストグラムを特徴量とし、AdaBoost<sup>1)</sup>を用いて色評価を行なうロジックを構築した。現在、定量的に十分な評価は行なっていないが、Fig. 9に示すように、鮮やかな色合いの風景画像などを良好に評価することに成功している。



Fig. 9 Frame evaluation in terms of color or colorfulness.

#### 4.8 総合評価

4.2節～4.7節で開発した各モジュールの結果を用い

て、それぞれの画像の自動評価を行なうのが「総合評価」モジュールである。

4.2節～4.7節の各モジュールで求められたそれぞれの評価値に、最適化された重みをつけて総合評価値を算出する。プリントに利用する画像など、利用に好適な画像を選択する際には、総合評価値が最高となる画像1枚を選択した後、当該画像が含まれる同一イベントおよび同一類似画像グループの画像の評価値を調整し、その後、総合評価値が残りの画像中で最高となる画像1枚を選択する、という処理を繰り返すことで利用に好適な複数枚の画像を選択する。

#### 4.9 iAgent制御

iAgentは、画像をハンドリングするアプリケーションから呼び出されるライブラリの形態をとる。アプリケーションと4.2節～4.8節の個々のモジュールの間を取り持つのが「iAgent制御」モジュールである。

「iAgent制御」には「画像読み書き」「処理結果入出力」の機能も含まれる。「画像読み書き」では、処理の高速化を図るため、評価対象となるJPEG画像を、4.3節～4.7節の各モジュールで必要とする解像度(VGA)に縮小処理を行ないながら高速にデコードするロジックを開発した。これにより、3M画素のJPEG画像を0.1秒/枚(Pentium4, 2.4GHz時)で読み込むことを可能としている。

### 5. iAgentの効果と課題

われわれは、4章で示したようにiAgentを構築し、その効果を確認するためにiAgentの自動選択に関するユーザ満足度調査を行なった。ただし、このユーザ満足度調査時点においては、4.7節の色評価モジュールは組み込んでいない。調査の概要は以下の通りである。

- 1) DSCユーザに直近撮影の300枚の画像を提供いただき、事前にその300枚中、気に入っている上位100枚を選んでいただく
- 2) 提供いただいた300枚をiAgentに入力し、iAgentによる評価値上位100枚までと、ユーザがあらかじめ選んだ100枚との正解率を求める
- 3) ユーザにiAgentによる自動選択100枚の結果を見せ、5段階評価で結果に対する満足度を回答いただく(有効回答19名)

その結果、正解率はユーザごとにかなりばらつきがあり、25%～75%となったが、想定した典型的なヤングファミリーユーザでは正解率も60%を超え、非常に効果的であることがわかった。一方、正解率が30%を割るユーザも存在したが、これらのユーザは主要被写体がペットや趣味の収集物であり、想定した典型的ユーザではなかったためと考えられる。平均正解率は45%と比較的低かったが、選択結果の満足度については、Table 1に示すように6割以上のユーザが「まあ満足」と

Table 1 Result of User Satisfaction with “iAgent” Frame Selection.

| 評価          | 人数 | 構成比(%) |
|-------------|----|--------|
| 1(非常に満足)    | 0  | 0      |
| 2(まあ満足)     | 12 | 63     |
| 3(どちらでもない)  | 3  | 16     |
| 4(あまり満足でない) | 4  | 21     |
| 5(非常に不満足)   | 0  | 0      |
| 合計          | 19 | 100    |

回答し、本技術の有用性が高いことを確認できた。また、正解率が60%を超えているユーザはすべて「まあ満足」と回答していることから、正解率60%が今後超えるべき基準の一つと考えられる。

次に、ユーザから iAgent の選択結果に対する不満点をヒアリングした結果を Table 2 に示す。Table 2 の (1) については、現行の「ボケ・ブレ評価」モジュールは顔が検出された画像のみを対象としていることが原因と考えられ、顔がない画像にも対象を拡大することで解決可能である。(2) については、いずれの例も類似画像分類自体は成功していたことから、総合評価モジュールでの同一類似画像グループ内の評価値修正方法を改良することで対応可能である。(3) については、ユーザ満足度調査後に4.7節の色評価モジュールを追加しており、これによる改善が期待できる。

Table 2 Main Causes for User Dissatisfaction with “iAgent” Frame Selection.

| 項番  | 不満な項目(複数名から回答があったもの) | 人数 |
|-----|----------------------|----|
| (1) | ボケた画像が選択される          | 9  |
| (2) | 類似画像から複数枚選択される       | 4  |
| (3) | 人物画像以外がほとんど選択されない    | 3  |

## 6. iAgent の適用先

iAgent は、一般ユーザ（現状ではヤングファミリー層向けに最適化されている）が撮影した大量画像の評価・整理・選択に効果を発揮する。そのため、PC や AV 家電向けの画像閲覧/整理/加工/アルバム作成アプリケーションや、店頭プリント端末向けのプリント画像選択支援機能などへの応用が期待でき、実際に製品化への検討が開始されている。また、良好な画像を選択するという用途のみならず、ボケ・ブレや露光不足などの不適画像に対して警告または自動削除を行なう用途などにも応用することができる。現状では PC (Pentium4, 2.4GHz 時) において、画像入力時に 0.5 秒/枚程度の処理時間がかかるが、演算量の削減や搭載モジュールの吟味などによって高速化を達成できれば、DSC などの機器への適用も視野に入ってくる。

今後、PC 向けの画像アプリへの搭載、店頭プリント端末アプリへの搭載から適用の検討を行ない、徐々に対象を拡大する予定である。

また、パラメータをチューニングすることで、一般ユーザ向けだけでなく、プロ用や業務用など、さらに幅広い用途に応用可能である。

## 7. まとめ

大量のデジタル画像の利活用を図る目的で、当初、実現は不可能と思われていたデジタル画像の価値の自動評価技術 “iAgent” を開発することに成功した。初期ターゲットをヤングファミリー層に設定し、6割以上の対象ユーザに受容される自動選択結果を得ることができた。この技術をうまく応用して商品化を実現すれば、デジタル画像の利活用を促進し、フォトライフを変えるポテンシャルを持っている。

今後の課題としては、特に不満の大きい「ボケ・ブレ」モジュールの改良のほか、ニーズの大きい「表情評価」「被写体人物評価」のモジュール追加開発、および処理速度の高速化、が挙げられる。また、上記技術課題の検討とともに、PC 向け画像アプリなどへの商品化検討や、ほかの用途へ向けたチューニング方法の確立を進めていく予定である。

## 参考文献

- 1) R. E. Schapire; Y. Singer. Improved Boosting Algorithms using confidence-rated predictions. Proceedings of the Eleventh Annual Conference on Computational Learning Theory. 80-91 (1998).
- 2) H. Nakamura. Digital Image Processing for the Frontier350 Digital Minilab. IS&T PICS 1999. 96-99 (1999).
- 3) 竹本文人, 依田章. 高品質を支える画像処理技術 “Image Intelligence”. 富士フィルム研究報告. No.49, 49-54 (2004).