

SlideChecker:プレゼンテーション資料の基礎的な定量的自動評価手法

SlideChecker: An Automated Basic Quantitative Evaluation of Presentation Materials

栗原 一貴 加藤 公一 大浦 弘樹*

Summary. 本論文ではプレゼンテーション資料のテキスト情報について、その基礎的な良し悪しを定量的に評価する手法、SlideChecker を提案する。SlideChecker は、プレゼンテーションツールを用いて作成されたスライド画像を入力として、テキストの分量、フォントサイズ、前景色と背景色の色合いについて診断を行う。272 枚のスライド資料に本手法による診断を適用し、有効に機能していることを示す。さらに、資料の分析評価結果を発表準備中および発表中に有効活用するアプリケーションを例示し、プレゼンテーションツールの新しい可能性を示す。

1 はじめに

現代社会においてコンピュータを用いたプレゼンテーションは一般的なものになりつつある。多くの人々がプレゼンテーションの機会を得るようになってきており、彼らのプレゼンテーションスキル習得を支援することは社会的課題の一つであると言える。それに応じる形で、プレゼンテーション術を示した技術書、ウェブサイトも多く出版・公開されてきており [1]、また学校教育においてもプレゼンテーションスキル学習がカリキュラムに組み込まれることが珍しくなくなってきている [13]。

このような教育的な解決方法に比べ、技術的な解決方法によってスキル習得を支援したり、スキル不要なプレゼンテーションを可能にするような取り組みについてはまだ発展途上であると言える。

プレゼンテーションスキルは、事前準備であるプレゼンテーション資料（ビジュアル資料）の構成に関するスキルと、発表時のプレゼンテーションデリバリースキルに大別される。後者について、栗原ら [10] は発表リハーサルを支援し基礎的なデリバリースキル改善を促す「プレゼン先生」システムを提案した。一方で前者について、標準的に用いられている Microsoft PowerPoint, Apple KeyNote などのオーサリングツールは、スペルチェッカーやテンプレートと呼ばれる発表構成のアウトラインを示すことでユーザを支援することはできるが、その多様な資料編集機能を用いた完成品が結果としてどのように聴衆に影響を与えるかについて、十分な情報を提供できていない。

そこで我々は、情報処理技術を用いてプレゼンテーション資料の良し悪しを自動的に評価する手法、SlideChecker を開発し、資料作成時・発表時それぞれにおいて発表者を支援することを目指す。当然のことながら、プレゼンテーション資料の総合的な評

価を完全に機械が判断することは難しい。本研究は、その総合的な評価を構成する多様な評価軸のうち、テキスト情報に注目し、発表者が意図せず行ってしまう結果として聴衆に不利益となるようないくつかの項目を排除することで、発表の質の底上げを図るものである。

プレゼンテーション資料のテキスト情報については、発表時間に見合うテキスト量であること、テキストのフォントサイズを小さくしすぎないこと、テキストの前景色と背景色として見やすい組み合わせを選ぶこと、などが好ましい条件として挙げられている。これらは、複雑な自然言語処理を導入するまでもなく現実的な性能を達成可能である。これらの情報を資料作成時および発表時にユーザに適切にフィードバックし、好ましくない状態にあることを警告したり、自動的な解決を行うか、もしくは半自動的な解決を支援する（図 1）。

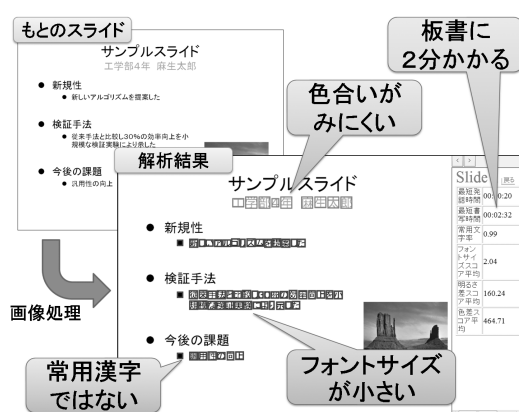


図 1. SlideChecker の活用例

2 関連研究

プレゼンテーションに関して、鈴木ら [12] は教育工学的な立場から、その構成過程の解明を明らかにすることを試みている。それによれば、発表者は資

Copyright is held by the author(s).

* Kazutaka Kurihara, 産業技術総合研究所, Kimikazu Kato, 日本ユニシス株式会社, Hiroki Oura, 東京大学 大学総合教育研究センター

料準備中に、聴衆の反応を想定しつつ内容を構成しているという。このような仮想聴衆をモデル化する上で、対話のモデル化という複雑な作業は発表者自身に委ね、「聴衆の視覚にどのように資料が映り、情報処理されるか」という基礎的なモデル化について、本研究では機械による自動的な支援を試みる。

Drucker らは画像処理などの技術を用いて複数のプレゼンテーション資料の類似度を計算し比較を可能にした [7]。テキストの色合いによるユーザの「見やすさ」を扱ったものとして、W3C はアクセシビリティの項目で、web コンテンツ中のテキストの前景色と背景色の望ましい組み合わせについての提言を行っている [2]。また、視覚障害者向けの web アクセシビリティを扱った福田ら [15] の研究や、印刷物と電子媒体における文字情報の読みやすさの比較を行った [11] などがある。我々はこれらの知見をふまえ、プレゼンテーション資料の色合いの分析に応用する。

発表中において、準備段階では想定されなかった状態に対応すべく、発表者の即興の表現力向上を扱った研究として [14][16] などが挙げられる。本研究では [14] のインタフェースをシステムが操作することで、資料中のテキスト情報に関する不具合の可視化や解決を行う事例を実装により示す。

3 プレゼンテーション資料のテキスト情報の評価指標とその算出方法

本章では SlideChecker において取り扱う、プレゼンテーション資料のテキスト情報の基本的な良し悪しを判定する 3 つの評価指標について、その算出方法を述べる。

プレゼンテーション資料の自動評価を実装する上では、プレゼンテーションオーサリングツールのアドイン、プラグインなどとしてソフトウェアを開発したり、オーサリングツールの出力ファイルを直接解析する手法などが考えられる。我々が提案する評価指標は、資料中の全てのテキスト情報について、(1) 文字ごとの外接矩形座標、(2) 外接矩形内部のビットマップ画像、(3) 対応する文字コード の 3 種類の情報が必要であり、その取得の方法は任意である。本研究では実装の基盤として [14] を用い、上記の情報の取得にパナソニック活字認識ライブラリ [5] を用いることで、様々なオーサリングツールにより出力された資料画像のみから評価が可能なシステムを構築した。以後はこれら 3 種類の情報が得られているものとして議論を進める。

3.1 テキストのフォントサイズ

いかに有益な情報であろうとも、それが聴衆にとって可読なサイズで表示されなければ効果はない。適切なフォントサイズ策定の具体案として、厚生労働省が平成 14 年に策定した「VDT 作業における労働衛生管理のためのガイドライン」[6] では、VDT(Visual Display Terminals) 作業者の心身の負担をより軽減のために、以下の条件が提示されて

いる。

- ディスプレイまでの距離は $l_{min} = 0.4[m]$ 以上
- フォントサイズは $f_{min} = 0.003[m]$ 以上

即ち、画面上に表示されるフォントの視野角を $\theta[rad]$ とするとき、

$$\tan \frac{\theta}{2} = \frac{f}{2l} \geq \frac{f_{min}}{2l_{min}} \quad (1)$$

が成り立つことが望まれる (図 2)。これを応用し、ディスプレイサイズやディスプレイまでの距離が変化しても、上記の条件を満たすことが、プレゼンテーション資料中の文字のフォントサイズにおいても必要な条件であると仮定すると、それは以下の不等式が成立することと同値である。

$$score = \frac{H_m f_p l_{min}}{H_p l f_{min}} \geq 1 \quad (2)$$

ここで、 H_m はディスプレイの高さ [m]、 H_p はディスプレイの縦解像度 [pixel]、 f_p はフォントの縦画素数 [pixel] であり、 $H_m : H_p = f : f_p$ が成立している。score をフォントサイズスコアと定義し、1 以上であることを推奨することとする。「(1) 文字ごとの外接矩形座標」から得られる f_p と、プレゼンテーション環境における H_m 、 H_p 、 l を上式に代入することで、その文字のフォントサイズスコアが得られる。たとえば、ディスプレイ高さ $0.2[m]$ 、ディスプレイ距離 $0.4[m]$ 、フォント高さ $12[pixel]$ 、画面縦解像度 $800[pixel]$ とすると、フォントサイズスコアは 1 となる。

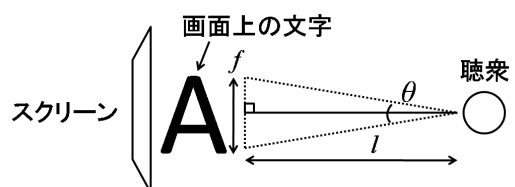


図 2. 聴衆と画面上の文字との関係

3.2 テキストの量

プレゼンテーション資料における情報の適切な量については、経験的に「1 スライド 1 分程度の分量」などと言われることがある [1]。これは、スライドの枚数から発表に必要な時間を見積もる作業を容易にする効果がある制約である。一方で一つのスライド内に含める情報量を常に一定量にとどめることは難しい場合が多い。そこでスライド中のテキスト量から、その発表に必要な時間の推定量を算出することができれば、資料を構成する上で有意義であろう。

本研究では発表に必要な時間の推定量として、「最小必要発話時間」および「最小必要書写時間」を算出する。前者は発表者がスライド内の全テキストを読み上げる際に標準的に消費する時間を、また後者

は聴衆がスライド内の全テキストをノートに休みなく書き取る際に標準的に消費する時間の下限を表す。これらはそれぞれ、ビジネス用か授業用かなど、プレゼンテーションの用途に応じて参考にできる。以下にそれぞれの算出方法を示す。

3.2.1 最小必要発話時間

栗原らのプレゼン先生システム [10] を用いて、日本語話し言葉コーパス [8] の講演音声について、印象評定タグで「ふつう」の速さにタグ付けされた講演音声 318 件を解析し、平均的な話速度 $4.95[\text{mora/sec}]$ を得た。プレゼンテーション資料の「(3) 対応する文字コード」情報から文字列を取得し、形態素解析エンジン (Yahoo 日本語形態素解析 [3]) により読み (ひらがな) に変換し、それを 4.95 で割ることで、標準的な発話時間を算出した。数式や英語アルファベットについては、計算から除外した。

3.2.2 最小必要書写時間

栗原らの音声ペン [9] では、日本人が日本語をペンで書く際にかかる時間 $T[\text{秒}]$ について、対象となる文字列の総画数を N とするとき、

$$T = 0.32N + 0.0831 \quad (3)$$

なる関係式を導出している。資料中の文字列からその総画数を得て、上式により平均的にかかる書写時間を算出する。また、その際に各文字が常用漢字かどうか判定し、非常用漢字についてはそれをユーザにフィードバックする。

3.2.3 テキストの前景色と背景色の色合い

プレゼンテーション資料中では、キーワードを強調するために文字色 (前景色) を多彩に変化させたり、資料の背景に多様なパターン画像や写真などを用いる場合があるが、意図せず情報を読みとりにくくしている場合も少なくない。

W3C では、Web コンテンツのアクセシビリティの項目において、すべての Web コンテンツにおけるテキストの前景色と背景色の組み合わせが以下の条件を満たすことを推奨している [2]。

コントラスト: 前景色と背景色 (それぞれ RGB255 段階で表現される) について、以下の式で計算される値の差をとったもの。

$$0.299 \times R + 0.587 \times G + 0.114 \times B \quad (4)$$

W3C では、その差を 125 以上とすることを定めている。

色差: 同様に前景色 (R_1, G_1, B_1) と背景色 (R_2, G_2, B_2) について、以下に式で得られるもの。

$$\begin{aligned} & \max(R_1, R_2) - \min(R_1, R_2) \\ & + \max(G_1, G_2) - \min(G_1, G_2) \\ & + \max(B_1, B_2) - \min(B_1, B_2) \end{aligned} \quad (5)$$

W3C では、この式の値を 500 以上とすることを定めている。

W3C の想定する Web コンテンツのテキスト色合い評価手法は、前提として HTML などにより明示的に前景色と背景色を知ることができるため、その計算は容易である。一方で現実には画面にレンダリングされるテキストの色合いは縁取り¹やスムージングなどの存在や背景に写真を用いた場合などにより影響を受けるため、その適用には限界もあると考えられる。

我々はこのような現状を踏まえ、W3C による Web コンテンツ上のテキスト色合い評価手法に若干の修正を加え、プレゼンテーション資料のテキスト色合い分析に応用する。提案手法は、明示的に前景色と背景色を知らなくとも、出力画像のみによる分析が可能である。

分析には、クラスター分析 [4] を用いる。クラスター分析は、特徴量ベクトルの集合を距離の近さに基づき段階的にグループ化することが可能である。図 3 のように、最終的に単一のグループになるまで繰り返しグループの併合が行われる。

今、資料中の 1 文字を表す「(2) 外接矩形内部のビットマップ画像」の処理を考える。画像中のそれぞれの画素の RGB 値 (0 以上 255 以下) を特徴量とする 3 次元ベクトルを作成する。リンケージアルゴリズムとしてセントロイド法を採用し、距離関数として式 4 (コントラスト) もしくは式 5 (色差) を用いる。

この条件でクラスター分析を行うことで、2 画素間のコントラストもしくは色差が小さい順に画素がグループ化されていく。もしも着目している文字が単一の前景色と単一の背景色のみからなる矩形に収まっている場合は、最終的にひとつのグループに併合される際の距離 (最終合併距離と定義する) が W3C 定義による前景色と背景色との間のコントラスト、色差に一致する。この最終合併距離を最終的な分析結果とすると、これは、本手法が「その領域での 2 大勢力グループのセントロイドを前景色、背景色とみだてて W3C 定義の手法で計算する」ことであることを意味している。背景が写真である場合など、背景色が乱雑で一定していない場合は、前景色のグループが顕在化せず、乱立する背景色の小グループへと埋没し、背景色の乱雑さに応じて最終合併距離が小さくなる傾向を持つ (図 3)。

提案手法の基礎的な性能を以下に述べる。PowerPoint を用いて単純な 8 色 (黒, 青, 赤, マゼンダ, 緑, シアン, 黄, 白) から背景色と前景色を選び (${}_8P_2 = 56$ 通り), MS P ゴシック 18 ポイントを用いてテキスト「あ」を描画した。その際に提案手法で得られるコントラストと色差を、式 4 および式 5 で得られる理論値で割った比率に変換し、その平均値と標準偏差を表 1 に示す。コントラストでは平均的には理論値と同様の値だがばらつきが大きい。一方色

¹ W3C には縁取りは前景色もしくは背景色として代用してよいという記述がある。

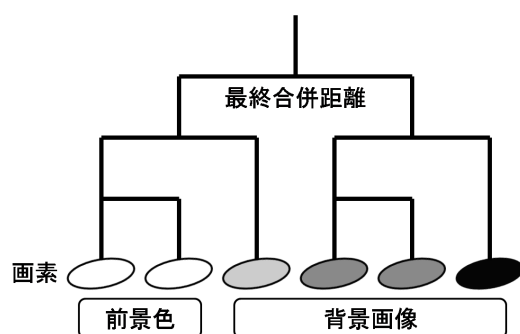


図 3. 1 色の前景色と乱雑な背景画像の場合のクラスタリング

表 1. 理論値との比較

	コントラスト	色差
理論値との比の平均	0.9846	0.9189
理論値との比の標準偏差	0.2671	0.06569

差は理論値から 1 割程度減衰する傾向にあり、そのばらつきは小さく安定している。これを踏まえて以後の分析では、テキストの色合い評価尺度として色差によるものを採用し、その推奨値を W3C 推奨値の 500 に、理論値からの減衰分と 2σ 分のばらつきの影響を加味し、 $500 \times (0.9189 - 0.06569 \times 2) \simeq 394$ 以上と定めることとする。より厳密には、様々なフォントにおける減衰の違いを考慮する必要があるが、今後の課題とする。

4 診断事例

我々は提案した SlideChecker を、中規模なデータに対し適用し診断した。以下にその事例の詳細を示す。

4.1 方法

学生を中心とした勉強会「CHI 勉強会」では、毎年国際学会 ACM SIGCHI における全ての登壇発表論文 (Paper と Note) について、1 論文につき 1 枚のスライドに内容を要約したものを作成し、それぞれ 1 分程度で説明を行う作業を繰り返している。2009 年 5 月 27 日に開催された 2009 年度の勉強会では、35 人がスライド作成と発表を分担した。全論文数は 272 であった。会場のプロジェクタ用スクリーンの縦サイズは 1.5[m]、最後列の聴衆までの距離は 10.88[m] であった²。我々は勉強会で作成された合計 272 枚のスライドについて、提案システムに

² 実際は奥行きのある会場であったため、会場正面に加えて会場中央にもう一台スクリーンを用意し、会場中央以後の聴衆用に発表画面を提示した。ここではフォントサイズ分析を簡素化するため、会場正面のスクリーンをもとにスクリーンサイズを計算し、会場中央スクリーン直前の聴衆を最後列とみなしてスクリーン距離を計算する。

より診断を行った。評価項目は、スライドごとの最小必要発話時間、平均フォントサイズスコア、および平均色差の 3 つである。

4.2 結果

表 2 に基礎的な統計量を示す。最小必要発話時間については、90.52%の発表で 1 分を下回っており、少なくともスライドの本文を読み上げるだけで 1 論文 1 分という基準を超えるケースはそれほど多くなかったと言える。しかしフォントサイズスコアについては、わずか 1 枚のスライドを除いて最後列の聴衆にとって基準値の 1 未満の値であったことがわかった。また色差については、全体の 31.48%のスライドで基準値の 394 未満であることがわかった。これらの項目については、基準を満たすようなひな型ファイルを事前に資料作成者に配布するなどして対策を図る必要があるかもしれない。

図 5 から図 9 までに、それぞれ最小必要発話時間の大きい上位 3 件、およびフォントサイズスコアと色差について値が大きい上位 3 件および値が小さい上位 3 件のスライドを列挙する。これを見ると、最小必要発話時間が大きいものは確かに文字数が多く、フォントサイズスコアが小さいものは確かに文字サイズが小さく、色差が小さいものは確かに文字がみにくくなっている傾向が直観的に理解できる。

表 2. 基礎的な統計量

	最小必要 発話時間	フォント サイズ スコア	色差
平均	34.63[秒]	0.4729	417.0
標準偏差	15.98[秒]	0.1293	101.2
基準適合率	90.52%	0.3704%	68.52%

5 アプリケーションの提案

我々が提案した SlideChecker は、前章で述べたように多数の資料の間での比較・ソートを可能にするだけではない。資料作成時、発表時および非発表者による診断時において様々なアプリケーションが実現可能であり、これらはオープンソースプレゼンテーションプラットフォーム「ことだま [14]」上で実装されている³。

5.1 資料準備時の支援アプリケーション

図 1 に示されているように、読み込まれた任意の資料画像列に対して資料診断機能を実行すると、フォントサイズ分析、および色差 (色合い) 分析により所定の基準値を下回った文字がハイライト表示され、改善が促される。また、常用漢字ではない漢字がハイライト表示される。

図 1 右端のサブウィンドウには、テキスト量分析により、資料全体および各スライドにおける最小必要発話時間、最小必要書写時間等が表示され、所定

³ 実際の動作についてはデモ映像も参照のこと

の発表時間に収まる内容を構成する上での参考にできる。

資料診断とは独立して、常にオーサリングツール画面上に「最小可読フォントサイズ」をアイコンで示し、ユーザがそのサイズを具体的に知ることができる機能も実装されている。そもそも入力するフォントサイズを最小可読フォントサイズ以上に制約する機能も考えられるが、実装に用いたことだけは ZUI を採用している。よってズーム状況によって見た目のフォントサイズが動的に変化してしまうため、このような仕様になっている。

5.2 発表時の発表者支援アプリケーション

テキスト量分析により得られた最小必要発話時間を用いてスライドを自動的に遷移することにより、実際にリハーサルを行い所要時間を記録することなく自動のループ再生が可能である。本機能はデモ会場における無人の発表などに有効である。

また、教育現場等での運用を想定し、最小必要書写時間までの時間経過を常に全てのスライドで計測および可視化し、聴衆に必要なノート取りの時間を確保することを発表者に促すことが可能である。

さらに、資料作成時に会場のスクリーンサイズや聴衆との距離が不明である場合に備え、発表当日に会場においてこれらの情報を入力することで、フォントサイズスコアが 1 以下である小さい文字が十分読めるサイズとなることを保証するよう、マウスポイント時に動的に拡大する機能を備えている (図 4)。

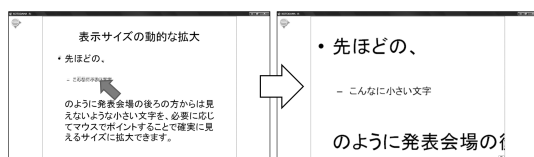


図 4. 発表者支援例：小さすぎるテキストがハイライトされており、マウスでポイントすることにより適切な大きさに拡大される。

5.3 発表時の非発表者診断アプリケーション

教育やビジネスにおける面接などで、発表者以外の人間により発表を評価する場合は実際問題として多い。この場合、全ての発表者の持参するコンピュータに本システムを組み込むことは、OS 環境の制約、発表時の計算負荷の制約から現実的ではない。しかし我々の提案手法は画像のみに基づいて行われるため、発表者 PC の映像出力をハードウェア的に分岐し取り込むことで発表者に制約を課することなく診断を行うことが可能である⁴。

6 まとめと今後の課題

本論文では、プレゼンテーションの資料中のテキスト情報について、テキスト量、フォントサイズ、

色合いについての診断を行うシステムを開発し、その運用例を報告した。また、診断を基にした有用なアプリケーション例を提案した。

今後は多言語対応、提案アプリケーションの有効性の検証、自然言語処理の導入による資料内容理解とそれに基づくユーザ支援、およびシステムを Web サービス化し公開することによる大規模な検証実験などを行っていきたい。また、今回はテキスト情報のみに注目したが、プレゼン先生システム [10] と組み合わせることによるマルチモーダルなプレゼンテーション評価も有望な方向性である。

参考文献

- [1] <http://www.isc.meiji.ac.jp/~re00104/ch12/>.
- [2] <http://www.w3.org/TR/AERT>.
- [3] Yahoo 日本語形態素解析. <http://developer.yahoo.co.jp/webapi/jlp/ma/v1/parse.html>.
- [4] クラスター分析. <http://aoki2.si.gunma-u.ac.jp/lecture/misc/clustan.html>.
- [5] パナソニック活字認識ライブラリ. <http://panasonic.biz/it/sol/ocr/sdk/>.
- [6] 新しい「VDT 作業における労働衛生管理のためのガイドライン」の策定について. <http://www.mhlw.go.jp/houdou/2002/04/h0405-4.html>.
- [7] Drucker et al. Comparing and managing multiple versions of slide presentations. In *UIST '06*, pp. 47–56, 2006.
- [8] Itou et al. A Japanese spontaneous speech corpus collected using automatically inferencing Wizard of OZ system. In *Journal of the Acoustical Society of Japan (E)*, Vol. 20, pp. 207–214, 1999.
- [9] Kurihara et al. Speech pen: predictive handwriting based on ambient multimodal recognition. In *CHI '06*, pp. 851–860, 2006.
- [10] Kurihara et al. Presentation sensei: a presentation training system using speech and image processing. In *ICMI '07*, pp. 358–365, 2007.
- [11] C. Timpany. The affect of type and colour on readability in print and screen based environments. In *ED-MEDIA '09*, pp. 797–806, 2009.
- [12] 鈴木 栄幸, 加藤 浩. プレゼンテーションの協同プランニング過程に関する事例研究. 電子情報通信学会技術研究報告. ET, 教育工学, 第 106 巻, pp. 47–52, 2006.
- [13] 岸 学, 小暮 敦子. 小学生のプレゼンテーション指導における評価項目の検討. 日本教育工学会第 23 回大会講演論文集, pp. 831–832, 2007.
- [14] 栗原ら. 編集と発表を電子ペンで統一的行うプレゼンテーションツールとその教育現場への応用. コンピュータソフトウェア, 第 23 巻, pp. 14–25, 2006.
- [15] 福田ら. Web アクアセシビリティ向上支援ツール: aDesigner. コンピュータソフトウェア, 第 23 巻, pp. 26–35, 2006.
- [16] 村田 雄一, 志築 文太郎, 田中 二郎. Shadowgraph: ペンの影を用いて OHP 風の指示ができるプレゼンテーションツール. WISS'08 論文集, pp. 73–78, 2008.

⁴ 本アプリケーションは、WISS 2009 において運用予定である。

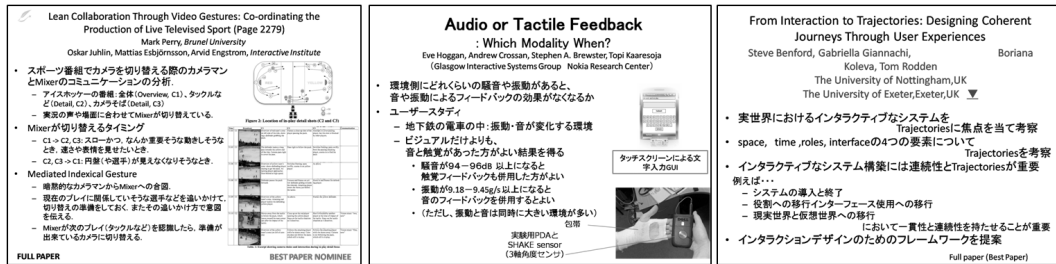


図 5. 最小必要発話時間の大きい上位 3 件: テキスト量が多い傾向がある .

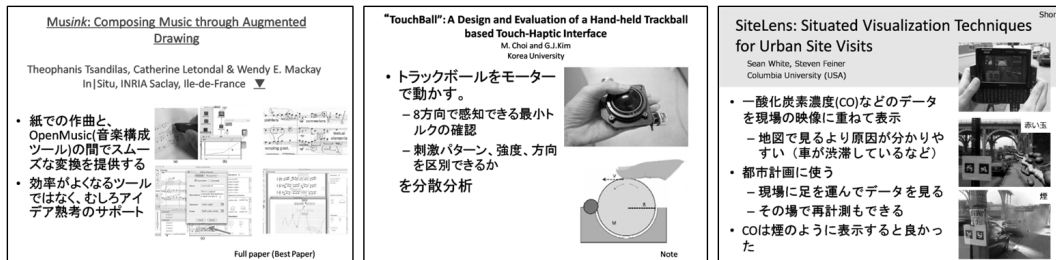


図 6. フォントサイズスコアの大きい上位 3 件: それぞれ本文のフォントサイズは 28pt, 32pt, 28pt である .



図 7. フォントサイズスコアの小さい上位 3 件: 引用した図中のフォントサイズが小さい傾向がある .

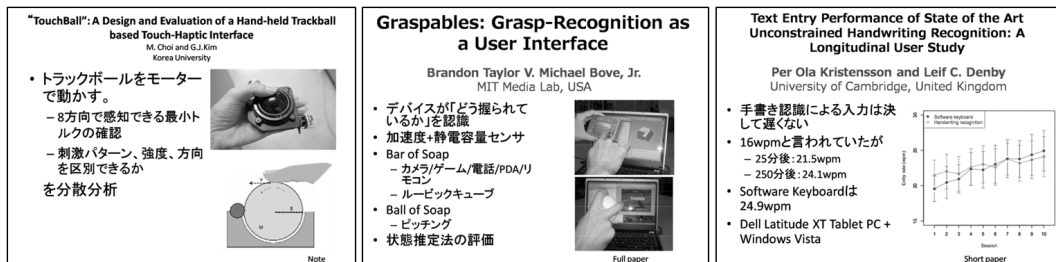


図 8. 色差の大きい上位 3 件: 白と黒を基調としており色差が大きい傾向がある .



図 9. 色差の小さい上位 3 件: 写真の上にテキストがあったり, テキストを含む引用画像の解像度が荒く滲んでおり, 見づらい傾向がある (著者により顔画像にはぼかしをかけてある .)