

車載認識装置における自転車認識技術の動向 トピック分析による認識技術の分類

○平川雅彦¹⁾

JFE テクノリサーチ (株)¹⁾

〒712-8511 岡山県倉敷市水島川崎通 1 丁目

Tel: 086-447-4286 FAX: 086-447-4285

E-mail: hirakawa@jfe-tec.co.jp

Trend of bicycle detection technology in driving support equipment Classification of detection technology by topic analysis

HIRAKAWA Masahiko¹⁾

JFE Techno-research Corporation¹⁾

1, Kawasakidori, Mizushima, Kurashiki Okayama 712-8511 Japan

Phone: +810-086-447- 4286 Fax: +81-086-447-2689

Email: hirakawa@jfe-tec.co.jp

【発表概要】

自動車に搭載する自転車認識技術の特許情報から解析する。特許情報の課題と解決手段を記述している文章を抽出し、トピック分析した。分析には LDA 法を用い特徴キーワード群を抽出する。信号灯、道路種別などの複合語は抽出できなく、このため、形態素3-gram の頻度で解析した。その結果、衝突の可能性などの特徴語を把握できるようになった。また、特許間の距離は潜在的意味解析により3次元の図形表示し、最も正確な位置関係のものを選定した。

今後は、解析精度をさらに向上できる手法を検討する。

【キーワード】

自転車認識, 特許情報, トピック分析, 形態素3-gram, 潜在的意味分析, ランキング分析, 対応分析, 多次元尺度法, R 言語

1.はじめに

最近、自動車に自動ブレーキなど障害物認識装置が搭載されるようになってきた。自動車アセスメントの安全審査基準でも自動ブレーキ搭載から自転車認識装置による安全確保が期待されるようになってきている。

そこで、特許情報から自動車車載装における自転車認識技術の動向をトピック分析を用いて調べることにした。

2.自転車認識技術

自転車を認識する技術は HYPAT-i2¹⁾ の概念検索を用いて抽出した。「自転車走行を認識する自動車に搭載したシステム」という概念で調査し、認識技術を始め 107 件の特許を得た。

この概念検索では特許公報の全文を対象として対話型処理によって重要度をランキング²⁾し、関連特許を選択する。選択した関連特許の情報も質問ベクトルに追加され再検索される。この適合性フィードバックにより検索精度は著しく向上している。

出願の推移では、1999 年から徐々に件数が増加しており、2006 年と 2015 年がピークとなっている。

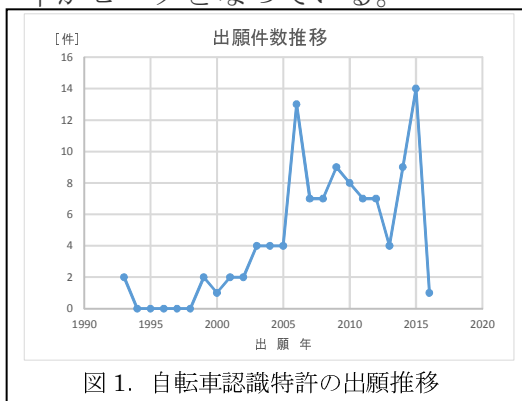


図 1. 自転車認識特許の出願推移

2006 年には認識技術がメインであり、早期開発の技術として運転支援がトヨタから出願されていた。2015 年にはデンソーからも支援装置が、また、

カタログ通板のイマージュが近接警報装置を 3 件も出願していた。

出願人では自動車部品のデンソー、自

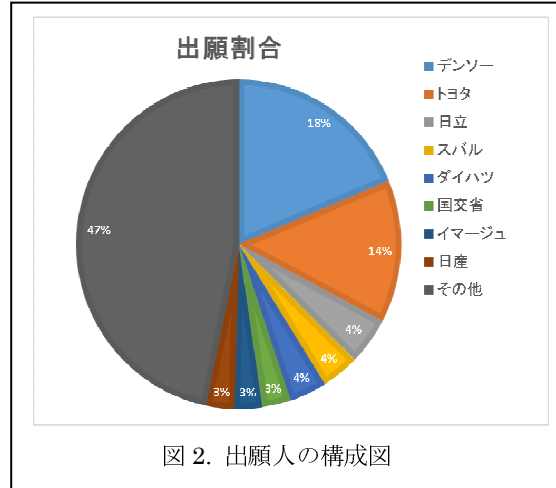


図 2. 出願人の構成図

動車メーカーのトヨタが多く出願しており、全体の 1/3 を占めていた。

解析は出願人の上位から 3 番目までを対象とし、デンソー、トヨタ、日立の 3 社とした。

3.動向分析

各出願人の特徴が出るようなトピックを抽出する。このため R 言語によるトピック分析: LDA 分析³⁾を行う。

3.1 トピック分析

代表 3 社の課題と解決手段を記載している文章を抽出し kw 頻度分析を実施した。

表 1 代表 3 社のキーワード頻度

No.	Friq.	特許数	KW	No.	Friq.	特許数	KW
1	1398	29	自転車	21	360	34	位置
2	1114	37	情報	22	328	29	出力
3	1107	35	判断	23	321	32	認識
4	942	39	車両	24	311	18	衝突
5	910	36	歩行者	25	311	10	ドライバ
6	693	37	検出	26	310	26	道路
7	631	36	ゾーン	27	308	7	標
8	597	19	物体	28	305	20	対象
9	570	38	存在	29	286	24	受信
10	568	38	自転車	30	281	7	危険度
11	531	13	移動体	31	279	29	距離
12	502	16	運転支援	32	267	22	送信
13	492	27	走行	33	265	30	コントロール
14	449	35	取得	34	261	15	E C U
15	421	25	画像	35	255	22	撮像
16	417	20	交差点	36	253	12	障害物
17	386	21	閾値	37	252	21	接近
18	373	32	設定	38	244	21	変化
19	373	10	対象物	39	243	12	警報
20	364	24	カメラ	40	233	20	運転者

運転支援、交差点、カメラなどのキーワードが特徴的に見えるが、このような短い単語だけでは特許公報の全体観が見えない。

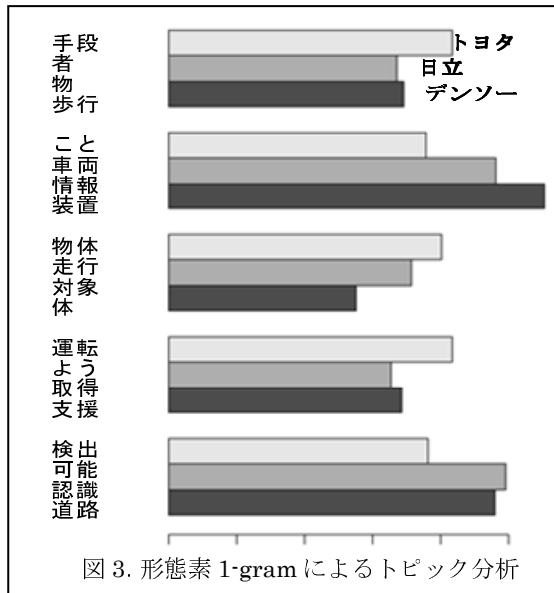


図 3. 形態素 1-gram によるトピック分析

LDA とは“Latent Dirichlet Allocation”の省略であり、文書中の単語の「トピック」を確率的に求める言語モデルである。

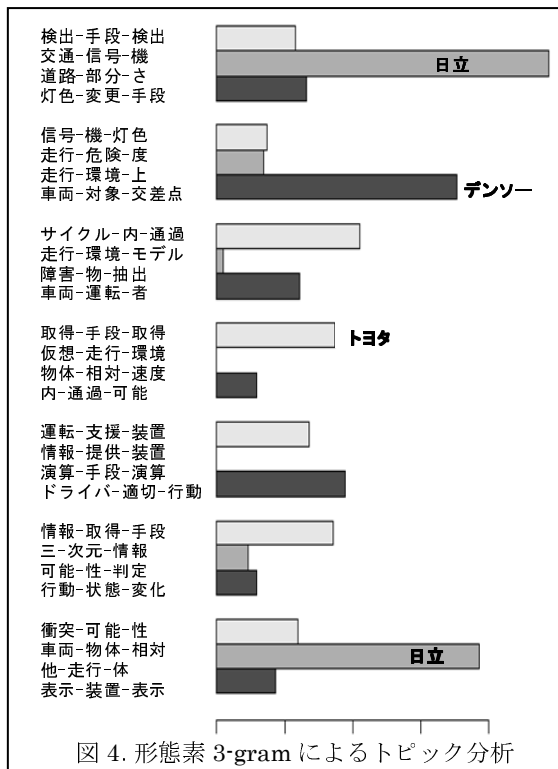


図 4. 形態素 3-gram によるトピック分析

自転車認識の特許公報に存在するトピックを抽出する。公報中の単語は単独に存在するのではなく、潜在的なトピックを持ち、同じトピックを持つ単語は同じ文章に出現しやすいと考えた。そこで、特許公報の全文を対象に、課題と解決する文章のみを抽出しトピック分析の対象とした。

具体的には R 言語 LDA パッケージを利用してトピックを抽出してみたが、各社の特徴を表現できるトピックは見つからなかった(図 3)。そこで、形態素 3-gram を用いてトピック内容が名詞の 3 単語になる様に前処理を行い対象用語の最低頻度を 5 に設定した。ここで、形態素 3-gram とは形態素が 3 個つながった組み合わせのことで、品詞は色々選択できるが、技術用語を抽出するため、名詞のみとした。

形態素 1-gram では日立に特徴があるトピックでも、検出/可能/認識/道路という名詞群ではトピックのイメージがつかめない。一方、形態素 3-gram⁴⁾では衝突-可能-性/車両-物体-相対と意味が明確になる。また、トピック kw の意味を解釈しやすくするために、灯色、二輪車などを形態素に分解する MeCab 辞書に追加登録した。

トピック数は 5 個と 7 個とを比較した。5 個では衝突の可能性に関するトピックが日立では際立っていることしか分からない。一方、7 個に増やすとトヨタが相対速度で、また、デンソーが走行危険度に特徴があることが判明した(図 4)。

3.2 トピックによる分類

トピックキーワードの単語群で特許の分類を行い、分類可能か確認した。トピック KW に着目し、このキーワードの同義語、または関連語を同じ色で着色した。また、これらのキーワードの

頻度を集計し、ランキング解析⁵⁾を実施した。ランキング最上位の頻度の20%~30%までを同じグループとした。

表 2. トピック KW による分類

No	KW1	KW2	KW3	KW4	頻度	Gr化	出願人
1	移動体	位置	自転車	接近	108		デンソー
2	移動体	自転車	検出	カメラ	71		日立
3	移動体	交差	軌道	対向車線	48	センサー	日立
4	検出	車両	距離データ	移動体	45		トヨタ
5	画像	立体物	移動	撮像	36		日立
10	運転支援	自転車	障害物	可能性	29		トヨタ
11	運転支援	依存度	運転者	プログラ	28		トヨタ
12	衝突回避	対象物	支援装置	ブレーキ	26	運転支援	トヨタ
13	障害物	運転支援	歩行者	検出	16		デンソー
19	色	灯	変更	交差点	121	交差点	デンソー
20	交差点	右折	横断	横断	22		トヨタ
24	物体	三次元情	情報	二次元	26		トヨタ
25	接近物体	車両	運転者	検出検知	23	認識装置	デンソー
26	物体	二輪車	物体検出	検出	17		デンソー
27	衝突	検出検知	物体	ドライバ	14		日立
33	運転	注意	ドライバ	プログラム	5		トヨタ
34	標	ドライバ	提示	情報	4	その他	デンソー

その結果、累積頻度による分類が自転車認識を代表するトピックとして、センサー、運転支援、交差点での処理、認識装置として分類できた(表 2)。

表 3. トピック分類の特許内容

No.	名称	出願人	Gr化
1	移動体位置検出システム	デンソー	センサー
2	車両制御装置、車両制御装置を搭載した車両、及び、移動体検出方法	日立	
3	車両制御装置	日立	
4	車両運動推定装置及び移動体検出装置	トヨタ	
5	立体物検出装置	日立	
10	運転支援装置	トヨタ	運転支援
11	運転者依存度判定装置及びプログラム並びに運転支援装置及びプログラム	トヨタ	
12	衝突回避支援装置	トヨタ	
13	運転支援システム及び車載用運転支援装置	デンソー	
19	衝突回避装置、及び走行支援システム	デンソー	交差点
20	走行支援装置	トヨタ	
24	物体検出装置	トヨタ	認識装置
25	車両用運転支援装置	デンソー	
26	物体検出装置	デンソー	
27	車両用衝突危険回避システム	日立	その他
33	運転支援装置、及びプログラム	トヨタ	
34	認知支援システム	デンソー	

センサーでは移動体検出という内容が多く、交差点での処理では情報提供に関するものがあり、その他にはプログラミングに関するものが特徴的である。また、日立がセンサーに関するものが多いことも特徴であった。

3.3 動向解析

次に、代表 3 社の用語の使い方の特徴を対応分析で調べた。分析には

KHCoder⁶⁾を使用した。トピック kw のなかでは「衝突」が日立に特有な kw で

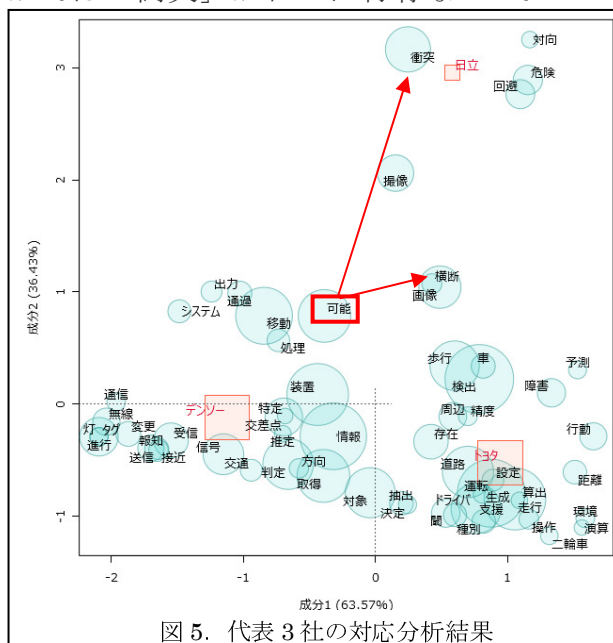


図 5. 代表 3 社の対応分析結果

あり特徴的である。「可能」は日立、デンソーで多く使われ、原文を調べると「衝突」と「横断」に関係している。トピック kw の衝突の可能性は日立に多く使われている特徴的用語になっていることが分かった。

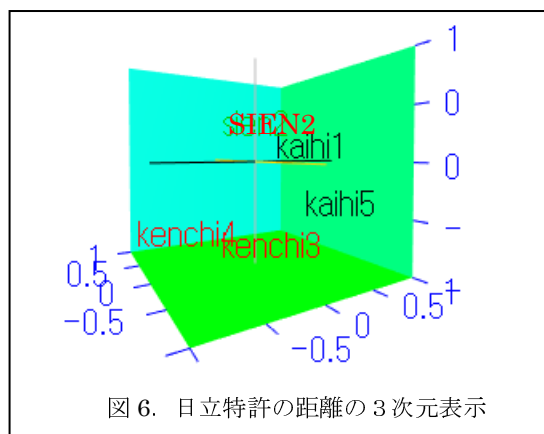


図 6. 日立特許の距離の 3 次元表示

日立特許の距離を潜在的意味解析により 3 次元表示する。rgl パッケージを使い 3 次元で回転できるように表現した。検知 3 と検知 4 が近傍に位置し、また、回避 1 と回避 5 も近傍に位置していることが分かる。回転角度に

よっては全体が重なるようになる。図を回転させ、それぞれの距離が正確に見えるように設定した(図6)。

支援2は表3の車両制御装置であり、センサーグループである。検知3、検知4も同じグループである。支援2が検知とかけ離れていることが良く分かる。

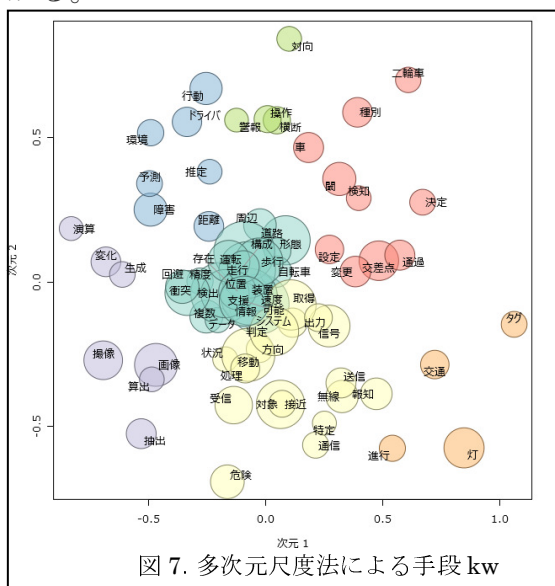


図7. 多次元尺度法による手段 kw

自動ブレーキでは運転者に報知することが主体⁷⁾であったが、自転車に警報を鳴らす手段も出てきた。危険を予測するという課題も出現している(図7)。

4. おわりに

トピック分析により特許公報の特徴を抽出でき、特に形態素 3-gram と組み合わせることによって特徴語を抽出できた。3連語の連続表示により特許の特徴がよく分かるが、どのトピック kw が関連しているのか対応分析などで確認する必要がある。

特許分類に関して、同一の概念を表すものをグループ化する意味的分類⁸⁾が今までは良く分類できたが、トピック kw による分類の方がこの意味的分類より効果的であり、適切に分類でき

ることが分かった。

今後、形態素 3-gram によるトピック分析の実績を積み、精度を向上できる分類方法を検討したい。

[謝辞]

本研究はアジア特許情報研究会の活動の一環として行った。原稿作成に際し、有効なご助言を研究会の各リーダー、並びに会員の方々から頂きました。ここに深く感謝申し上げます。

参考文献

- [1] HYPAT-i2.
https://www.hatsumeimei.co.jp/hypat_i2/function.html (参照 2017-9-22)
- [2] 真野博子ほか. “文献検索におけるランキング技術2” Ricoh Technical Report No. 29 p21-p30
- [3] 芹澤翠, 小林一郎. ”文章内のトピック数を考慮したトピック追跡の試み” 言語処理学会 第18回年次大会 発表論文集 p1196-p1199
- [4] 石田基広 “Rによるテキストマイニング入門” 森北出版(株) p75-p79
- [5] ぱっとマイニング JP.
<http://www.pat-mining.com/> (参照 2017-9-22)
- [6] KHCoder.
<http://khc.sourceforge.net/index.html> (参照 2017-9-22)
- [7] 平川雅彦, 安藤俊幸. “自動ブレーキの国内と中国の動向” 第12回情報プロフェッショナル シンポジウム p79-p84
- [8] 平川雅彦, 安藤俊幸. “自動ブレーキの周辺技術の動向” 第13回情報プロフェッショナル シンポジウム p83-p88