

移動・環境要因を用いた機械学習による猿の出現予測

メタデータ	言語: ja 出版者: 静岡大学 公開日: 2017-06-09 キーワード (Ja): キーワード (En): 作成者: 中井, 一文 メールアドレス: 所属:
URL	https://doi.org/10.14945/00010237

静岡大学大学院創造科学技術大学院情報科学専攻博士論文

移動・環境要因を用いた 機械学習による猿の出現予測

中井 一文 (5524-5017)

2016年12月

指導教員：杉浦 彰彦

目次

第1章	序論	3
1.1	本論の背景	3
1.2	本論の目的	4
1.3	本論の構成	5
第2章	本研究の原理	6
2.1	動物の行動調査	6
2.2	マルコフモデル	15
2.3	ベイズ推定	24
2.4	Support Vector Machine	34
2.5	本論の関連研究	46
2.5.1	RSSI を利用した進行方向の推定	46
2.5.2	RSSI を利用した位置推定	47
2.5.3	赤外線情報を利用した害獣の計測	47
2.5.4	猿の襲来に関する分析	47
2.6	実験システムの構成	48
2.6.1	システム設計	48
2.6.2	猿の行動収集	49
2.6.3	AP の稼働率	50
2.6.4	受信信号強度による接近検知	51
2.7	予測システムの構成	53
第3章	移動要因を用いた猿の出現予測	54
3.1	緒言	54
3.2	原理	55

3.2.1	マルコフモデル	55
3.2.2	猿の出現の定義	55
3.3	実験方法	58
3.4	出現/非出現の2クラス問題の結果	61
3.5	出現 AP のマルチクラス問題の結果	62
3.6	結言	64
第4章	環境要因を用いた猿の出現予測	65
4.1	緒言	65
4.2	原理	66
4.2.1	ベイズ推定による出現予測	66
4.2.2	SVM による出現予測	66
4.2.3	複合型分類器	68
4.3	猿の出現の定義の時間的分割	70
4.4	予測で使用するパラメータの分析	71
4.5	実験方法	73
4.6	実験結果	74
4.7	移動要因と環境要因を組み合わせた出現予測	77
4.7.1	移動要因と環境要因の組み合わせ予測手法	77
4.7.2	移動要因と環境要因の組み合わせ予測結果	77
4.8	結言	79
第5章	結論	80
5.1	本論のまとめ	80
5.2	今後の展望	82
	謝辞	83
	参考文献	84
	付録 A 使用機器の諸元	96

第1章 序論

1.1 本論の背景

現在，日本におけるニホンザル（以下は猿とする）による農作被害は年間十数億円であり，農林水産省をはじめとして，各地域の行政機関や農家から問題視されている．猿は学習能力が高く，単純な威嚇装置ではすぐに慣れてしまったり，電気柵も乗り越えてしまうなど，人が介在しない簡易な方法では追い払いづらい．人以外の方法としては，犬による追い立てもあるが，犬の行動範囲を農作地に限定するための柵が必要であるため露地栽培においては必ずしも適用できず，また，犬を飼うというコストが発生する．さらに，猿は人に慣れると人がいるときにも街中に出没するようになり [1]，住民にとっても危険な害獣となりえる．猿を人に慣れさせないためには，人の手により威嚇・追い払いを行うことで，人が怖い存在であることを認識させることが重要である．しかし，人が一日中農作地を監視するのは大変であり，現実的ではない．そのため，猿の動向を監視するシステムが必要である．

猿の動向を監視する既存のシステムとして，猿に発信機付きの首輪を取り付ける手法が存在する．発信機を取り付ける手法では，猿が農作地に近づいた際に発信機から送信された電波を農作地に設置した受信機が受信した場合にサイレンや回転灯で猿を追い払うことを一次的な対処とし，同時に猿の接近を地域住民に知らせることで人による追い払いを促している．しかしシステムの問題として，まずは猿を捕獲し，発信機付きの首輪を取り付ける必要がある．次の問題として，群れの中に発信機を付けた猿がない場合はシステムが動作しない．猿は群れが大きくなるとある程度分割した群れでも行動しており，分割した群れの中に発信機を付けた猿がないということも想定される．さらに，発信機を付けたとしても2年程度で電池が切れてしまい，猿を継続的に捕獲して発信機を付け直す必要がある．また，猿が接近し通報が行われた場合でも，住民の外出等で準備不足の場合は即時的な対処が難しくなる．

1.2 本論の目的

本研究では、猿の出現を予測するシステムを作成することを目的とする。

猿は餌を求めて農作地や山を移動していくが、この移動がモデル化できれば、猿の出現を高い確率で予測できる可能性がある。しかし、現在までに短時間の猿の移動をモデル化する研究はほとんど行われていない。猿の出現に関するモデルが天候などの環境パラメータや猿の目撃情報などの人間が観測しやすい変数に従属しているとき、猿に発信機をつける必要がなくなり、猿が次にどこの農作地に出現するかの予測をすることができる。猿の出現しやすい日時をあらかじめ地域住民に知らせておくことで、先立って追い払い器具を用意しておくなど準備不足問題の解決、ピンポイントでその日時だけ農作地を監視しておくなど時間の短縮による負担の軽減ができるようになる。

本論では、予測を行うために猿に発信機付きの首輪を取り付け、受信機により動向データを採取する。採取したデータから猿が農作地に出現したか、それとも非出現であったかのラベル付けを行う。

ラベル付けデータから猿が1日毎にどのように農作地や山を移動しているかをマルコフチェーンでモデル化し、マルコフの状態遷移確率を基に猿が出現するかどうかの予測を行う。

さらに、データを1時間毎に分割し、分割されたデータを気象データ等の環境パラメータと合わせて、ある時刻に猿が出現するかどうかの予測を行う。予測手法として、ベイズ推定による予測、SVM(Support Vector Machine)による予測の2種類について実験を行い、予測精度について比較・検討を行う。

最後に、移動要因による1日単位の予測と、環境要因による1時間単位の予測を組み合わせたものを本論の予測システムの提案手法とし、1時間単位の出現予測の精度を検証する。

1.3 本論の構成

本論では、移動・環境要因を用いた機械学習による猿の出現予測を行う。

本論文は以下の5章で構成されている。

本章“序論”では、本研究の背景と必要性、及び目的と本手法の有用性について述べた。

2章“本研究の原理”において、動物の行動調査、マルコフモデル、ベイズ推定、SVMの説明を通して、現在の研究分野における状況をまとめる。併せて本論に関連する研究についても説明を行う。また、本論の3章、4章で提案する猿の出現予測システムの説明を行い、システムの構築や猿の行動収集の実験も説明する。

3章“移動要因を用いた猿の出現予測”では、猿がどのように集落間を移動していったかという移動要因を用いた場合の予測システムの検証結果を示す。

4章“環境要因を用いた猿の出現予測”では、猿の動向を観測しなくても得られるパラメータとして環境要因を用い、環境要因と猿の出現との相関関係を示しながら猿の出現予測を行った結果を示す。4章では、本論の提案手法として移動要因と環境要因を組み合わせた予測手法についてもまとめる。

5章“結論”において本論のまとめと今後の展望について述べる。

第2章 本研究の原理

2.1 動物の行動調査

本節では、猿をはじめとした動物類の行動の先行研究について述べる。

猿は一般的に餌を求めて移動する。中川 [2] によると、猿の食性を年間の時間的な変化にわたって定量的に記載した研究はほとんどなかった。理由として、猿の生息地の地形が急峻であったり、植生が密であるため、年間に渡って量的データを得ることが困難であったためとしている。そこで、宮城県金華山島を調査地域として、猿の食性を数年にわたって調査した。調査報告では、季節の変化に従って猿の食性が変化していくことを、食物ごとの採食時間の割合で示している。

Tsuji ら [3] はさらに長期に渡って金華山の猿の食性の変化を調査している。調査では、食物の種別毎に猿の採食時間の割合を月毎の変化としてあらわしている。それに加えて、四季毎に食物の種別の採食時間の割合をみたとき、年によってどの食物に強く依存するかが変化することを示している。

深谷ら [4] は猿の食性の選択性について着目している。選択性とは、ある時期における食物資源に対して採食対象となるかどうかのことであり、食物資源の利用可能度と猿の採食量の対応をとることで、選択性に違いがあることを明らかにした。研究によると、特定の時期に選択度が高い種は年間を通して選択度が高く、猿の食物資源として重要であることを推察している。また、食物資源量で見たとときに単独種だけでなく、複数種が多く存在する場所でよりカロリーを摂取している可能性を指摘している。

江成ら [5] は猿の食性の季節的な変化に加え、農作物と野生種との採食割合についての調査を行っている。調査地域である白神山地北東部では、7-8月と11-12月に採食地として農地が選択されることが多いとしている。理由として、野生種の食物資源量が減少することや猿の栄養蓄積期が重なることや、農家の農閑期による見張りの低下などの複合的な要因が挙げられている。

吉田ら [6] は農作物に絞り被害品目などの詳細を調べている。季節により被害の割合は変化するものの、どの季節においても何かしらの品目の農作物被害がある。調査で興味深い点として、収穫されなかった農作物の残滓や生ゴミについても猿の餌として観察し、季節的な変化を捉えており、猿を誘引する原因とならないように残滓や生ゴミの適切な処理の必要性を訴えている。また、農作地から林

縁までの距離によって被害の割合が違うことを示している。農作地から林縁までの距離は季節によって変化しており、森林内で食物を得づらいう冬期は、人間と遭遇するリスクが高くなったとしても森林から離れた農作地まで猿が到来するようになる。また、逆説的に、森林内で食物を得やすい秋期などでは、林縁のそばの農作地でも十分な量の食物が得られるため、林縁から離れた農作地まで猿が来ないと考察している。

辻 [7] は総説として猿の食性をまとめており、これによると全国的に食性の季節変化は見られ、毎年の結実状況が違うことのひとつの原因として気象変動を挙げている。食性は、季節変動などの時間的変異だけでなく、高緯度・低緯度地域などの空間的変異、性・年齢・生理状態などの個体間変異もあるとしており、それぞれのデータを示している。辻は、猿が生息地の食物・物理環境や猿自身の社会的立場によって採食に関する行動を柔軟に調整できると述べており、食性の長期データの蓄積や地域間比較の充実が今後も必要であるとまとめている。

また、猿の生態系を調査する別の視点として、大井 [8] によるツキノワグマと猿の種間関係をまとめた報告がある。クマと猿の捕食-被食関係や、クマが猿を採食場で追い払ったなどの直接的な種間作用の事例はないとしているが、食物の消費競争ならば起こっている可能性があるとしている。報告の中で、猿は群れで行動するため、食物の探索に優れ、嗜好性の高い食物の獲得に失敗しても、幅広い代替食物でエネルギーを補うことができると述べている。

大井ら [9] は別の報告でクマの食性の変化を詳細に述べている。山の食物が少なくなるとクマの人里への出没が増えることを結びつけ、年毎の山の食物量の出来からクマが人里へ出没するかどうかの予測ができる可能性を示唆している。また、農作物やカキ・クリなどの果実が人里での主な食物であると言われていたが、農地の周りの草本の葉などもクマの重要な食物になっている可能性も指摘しており、理由については、農地周りの草本は栄養の含有量が比較的高いものが多く、さらに光環境がいいので現存量も多いことを挙げている。さらに農地周りの草本はクマ以外に猿の誘引の原因になっていることにも言及している。

齊藤ら [10] も農作地と森林の関係に注目している。研究における観察によると、猿が農作地で採食しているときも、実は大半の猿は農地に接する林に潜んでいることを確認している。また休息するのは雑木林や杉林の暗い下層部にいたとしている。実験では、猿の行動圏を制御する試みとして、耕作放棄地および水路の周囲を中心に雑木や下草を伐採したところ、猿がほとんど近づかなくなったと報告している。下草などの伐採によってオープンスペースができ、隠れる場所も食物もなくなったために猿の滞在時間が減った可能性に言及している。

猿の行動圏を調査した研究 [11] はあるが、猿が出現する可能性や農作物被害リスクの予想を立てた研究は少ない。江成ら [12] は猿の群れの分布や、山間部・平地などの地形的な特徴を加味し、リスクマップを作る研究を行っている。研究では、地形的に猿に適した場所を分析し、猿の群れの所在地も含めてアクセシビリティを計算し、それらから農作物のリスクマップを作成している。

テレメトリにより行動圏を調査した研究としては永田 [13] によるシカの研究がある。給餌場を人工的に用意し、給餌場の有無でシカの行動圏が変化するかどうかの実験を行った。実験の結果、給餌場を利用した個体は給餌場を中心に行動圏を形成するようになり、さらに定住型の行動圏を小規模に変化させていく様子が見られたとしている。すでに給餌場となっている地域に対して動物の行動圏を調べる研究はあっても、人工的に給餌場を用意して行動圏の変化を調べる研究は少なく、永田の実験結果が他の動物にも適用できるのであれば、農作地をはじめとした集落を動物の給餌場にさせないための努力は欠かすことができず、一度給餌場として行動圏を形成されてしまえば、追い払いなどの余分な労力が必要となることが想定できる。

石塚ら [14] は農作地をすでに給餌場として利用しているシカの行動圏の変化を、季節的・時間的に調べている。行動圏は季節的には変化しなかったとしているが、時間的な変化として昼は森林域、夜は水田に滞在していると報告している。また、行動圏のコアとなるエリアが農作地に隣接したエリアにあり、農作地への依存度が高いことを推察している。行動圏が季節的に変化しなかった理由として、イネの耕作時期以外にも水田の周りには畝畦の草本があり、これらを食物として利用していることを挙げており、こういった理由は大井ら [9] の報告とも一致するため、人里にあらわれる動物に共通する性質と考えられる。昼は森林域、夜は水田に滞在する理由として、昼間に人間に遭遇することを避けるためと考察しており、水田が採食場所、森林が休憩場所、水田と森林の間に存在するコアエリアが林縁部にあるというのは、採食場所と休憩場所との移動に使われているとしている。

猿の行動を予測する研究は少ないものの、分布変化を決定木で予測する研究は江成 [15] によって行われている。研究では、住宅地の割合、標高、植生、平均気温などをノードとして決定木を構成し、猿の群れが拡大・非拡大の傾向にあるかどうかの予測を行っている。判別基準が十分に抽出されていないために予測の判別精度は低いが、平均気温などのファクターは猿の群れの拡大・非拡大など影響要因になるとしている。判別精度が低い原因のひとつとして、猿の不在情報の収集の難しさを挙げている。これは、猿の群れの実態調査の不足等による情報欠損により、実際は群れがいる場合であっても正解データでは不在となっているなどの理由により、誤分類が増えたために推定精度が低下した可能性を指摘した考察である。また、この調査の一環で明らかになった事実のひとつとして、現在実施

されている捕獲事業は、増加傾向にある個体群に対して分布制限要因になりづらい傾向を述べている。

猿以外の動物の行動をモデル化した例として、有本ら [16] の研究がある。研究ではクマに GPS 首輪を付け行動を調査し、行動区分を「移動」・「活動中の滞在」・「休息中の滞在」としてベイズモデルを使いモデル化を行っている。行動区分のなかで「活動中の滞在」とされた場所を集中的に調査することで、植生や生活痕跡の記録に役立っている。調査結果から、クマが集落周辺を給餌場として利用していることを報告しており、また、集落周辺では夜間の活動が増えていることから、クマも人の活動を避けるように行動していることを考察している。

山端 [1] は集落ぐるみの追い払いによる猿の農作物被害軽減効果について報告している。報告によれば、猿は人に慣れてしまうと集落に出没するようになり、農作物の被害が加速度的に増加してしまう。そこで有効な対策手法としては集団による追い払いがあるが、集落の農家全員が猿を追い払うという行動は合意形成が困難であり、現在の実態としては個々の農家が自分の農地から猿を追い払うだけにとどまっているため、被害の軽減が難しい状況にある。山端は三重県内の 6 地区において集落ぐるみの追い払いを実施した。結果として、2 年間という短期間ながら 6 集落中 4 集落で農作物の被害軽減効果が見られた。また、効果の見られた集落においては猿が人に慣れている度合いも下がっており、猿にとって人間が脅威な存在であるという認識に変化している。これらの 4 集落では追い払いの実施率が高く、その理由としては、自治会組織に加えて農家組織が存在することで、集落ぐるみの追い払いを行うための合意形成が可能であったとまとめている。集落によっては 65 歳以上農家率が高かったり、昼間に被害対策に参加できるのが女性ばかりであったとしても、合意形成さえできていれば集団ぐるみの追い払いは可能であることを検証している。

また、山端 [17] は獣害対策の進展が農家の農地管理意識に及ぼす効果の検証も行っている。集団ぐるみの追い払いを実施する前の調査では、農業を辞めたいと思うかという設問に対して約 4 割が肯定的な回答をしており、農業を辞めたい原因の 2 番目にくるのが獣害であることを調べている。獣害の被害程度と獣害対策への否定的な意見・農地管理への否定的な意見は相関関係にあり、獣害被害が多いほど獣害対策や農地管理へのやる気をなくす。同様に、自分たちで行った獣害対策により被害が減れば減るほど獣害対策や農地管理を肯定的にとらえていくように変化していくと報告している。

さらに、中村ら [18] は加害群となる猿について、住民がどういう理解をしているかを調査している。加害群の猿が行動する地域において、加害群が利用する頻度の高さによって高リスク地域、中リスク地域、低リスク地域と分け、それぞれの地域で住民が猿の頭数を正しく把握しているか、また、住民が猿の頭数を何頭まで許容できるかをアンケート調査した。調査の結果によると、どの地域にお

いても実際の群れの頭数よりも少ない群れであると考えている住民が多いが、高リスク地域ほど猿の頭数が多いと感じているとしている。また、群れの頭数を何頭まで許容するかの結果では、中リスク地域、低リスク地域ではおおよそ半数程度の住民が0頭、高リスク地域においては75%の住民が0頭までしか許容しないとしている。頭数を実際よりも少ないと誤認している理由としては以下の理由が挙げられている。

1. 調査対象地域が住宅地であるために出没した全頭を確認するのが困難であること
2. 少数のハナレオスなどのグループを目撃することが多いこと
3. 出没する時間に出勤などしており、目撃する機会が少ないこと

また、許容する頭数が0頭であった理由としては、どのリスク地域においても深刻な被害状況による猿の存在への否定感を挙げている。

以上のように人の立場に立ったときの視点から被害軽減効果を心理的な変化として捉えるだけでなく、山端 [11] は集団ぐるみの追い払いが猿の行動域にどういった変化を与えるかという実験を行っている。実験では、同一猿群の行動域に存在する複数の集落で、集団ぐるみの追い払いを実施する集落と実施しない集落を設けた。その結果、農作物被害が減った集落は猿の行動圏から外れたことを定量的に示している。その一方で、追い払いを実施しなかった集落では被害が増加するところもあり、全ての集落で追い払いを行わなければ、猿の出没が増加してしまう可能性も示唆している。人里の農作物の栄養価が高いことが原因として猿の個体数が増えることが被害増加の原因ならば、集団ぐるみの追い払いを全ての集落で行うことで猿の個体数の抑制も可能ではないかと言及している。

追い払いの効果についての定量的な研究として、山田 [19] はいくつかの動物種が生存や繁殖に関わる行動の最適化を行っている例を取り上げながら、動物にとって危険な存在のひとつである人間と、猿の行動の最適化の関係を調べている。研究では猿について、危険に対する行動変化の指標のひとつとして使われている Flight initiation distance (FID: 対象動物に対する接近可能距離) を用い、1つの猿の群れが追い払い対策の異なる6集落においてどのようなFIDになっているかを報告している。報告によると、コドモとアカンボウ持ちのメスがメスやオスよりも長いFIDをとっていた。つまり、人に対しての警戒心が強い。また、まったく追い払い対策をしていない集落と対策をしている集落では有意な差があったとしている。

猿の追い払い手法のひとつとして追い払い犬(モンキードッグ)の利用がある。山口ら [20] は西日本における追い払い犬事業の成果について調査している。報告では、自治体担当者らからアンケート

をとり、地域全体で追い払い犬事業に積極的に取り組んでいる自治体と事業を中止した自治体を比較して事業の問題点を指摘している。報告によると、追い払い犬事業の効果として農作物被害が減少したことや人による追い払いよりも効果的であるとした結果が高い割合で挙げられる一方で、猿が別の場所に出没してしまう問題も挙げられている。また、追い払い犬事業が中止に至った自治体においては理由として、地域の協力体制が不十分であったり、科学的なデータがないために効果の有無がわからなかったことが挙げられている。山端 [1] の追い払いに関する報告には追い払い犬を用いた旨の記載はないが、地域の協力体制が必要であることを述べている点が山口ら [20] と共通しており非常に興味深い。

現在の追い払いは人や犬によるものが多く、光や音などを自動発生させる手法では猿がすぐに慣れてしまうため効果がほとんどないとされている。田中ら [21] は超音波音とシチメンチョウ音声に対する猿の忌避反応を定量的に調べている。研究では、動物に嫌悪感を与える可能性があると言われる超音波音と、猿の被害防止に効果があるとされているシチメンチョウ音声を猿の採食時間に再生し、忌避による行動の変化を報告している。超音波音を使った実験では、忌避効果は実験開始後 2~3 日しか維持されず、長期的にみれば採食行動の抑制に対してほとんど効果がないとしている。また、シチメンチョウ音声を使った実験では、新奇な生物的音刺激として強い逃走反応を引き起こすが、その効果はごく短時間で消失するとしている。研究のまとめとして、猿が逃走反応を示す音や接近を躊躇させる音でも、単純に音刺激を呈示するだけでは猿はすぐに馴化してしまうと述べている。

動物の行動を位置情報として調査する方法として、大きく分けて GPS を利用する方法と VHF を利用する方法の 2 つがある。宇野ら [22] は GPS 受信機の精度を調べており、針葉樹林や広葉樹林などの植生や、斜面や谷などの地形による精度の変化を報告している。GPS 受信機の製品によって精度の違いはあるものの、二次元座標の測量であれば、概ね良好な結果が得られるとしている。また、報告の中で GPS 受信機を鹿に取り付け、鹿の行動により GPS 受信機の向きが変化する状態においても、測量の結果に有意な差はないとしている。

佐伯ら [23] はラジオテレメトリの方法についてよくまとめている。報告によると、テレメトリの方法は調査対象とする動物や、調査項目によって選ぶ必要があり、効率良く代替できる方法があればそちらを使うべきであると述べている。首輪型が主流であるテレメトリ装置を動物に取り付けるにしても、動物のサンプリング数は研究目的に沿って予め把握しておくことが望ましく、また、取り付けのための捕獲は一定期間中に必ずしも必要数が得られるとも限らないとしている。テレメトリの方法である GPS と VHF を比較すると、GPS は位置情報が記録され続けていくので動物を追跡する必要は

ないが、首輪型の装置を回収するまではデータを知ることができず、首輪の回収の難しさや電池持ちの悪さもデメリットとして挙げられる。VHFを用いる方法では、電波強度から動物の位置を把握することになるので位置情報の精度を上げづらく、詳細な位置を知るためには動物を常に追跡しなければならない。本論では後述するが、猿を追跡するのではなく、猿が接近してくるのを検知できればよいため、VHFでのラジオテレメトリを使用している。

佐伯ら [24] は動物の個体追跡のデバイスとして、一般的なGPS搭載携帯電話端末の利用を検討している。すでに述べられているように、首輪型のGPS装置は確実な回収が難しい。また、別の携帯電話端末としてPHSがあるが、PHSは出力が弱いために基地局との通信距離が短いという点を活かしてある程度の位置推定ができるが、誤差は数十mから数kmともいわれており、GPSより誤差が大きい点が不利である。また、基地局は都市圏に多く設置され、山間や緑地では基地局は少ない傾向があるため、市街地であったとしても緑地に存在する個体の追跡には向かない。そこで3G通信網を利用することができる一般的なGPS搭載型携帯電話端末を動物の個体追跡に用いることができるか調べている。調査では以下の項目について触れている。

1. データ回収成功率
2. 利用可能エリア
3. 位置の誤差
4. 端末の大きさ・重量・形状
5. 電池寿命
6. 装着方法

調査によると、通信網のサービスエリア内ではほぼ確実に位置情報が取得でき、誤差も十数mから数十mで収まり、イヌやネコに装着した結果からも端末の大きさ・重量・形状に問題なく既存のVHF機器やGPS機器と大差はなく、動物の追跡デバイスとして使える可能性があるとしている。適用先としては、携帯電話の基地局と安定して通信できる都市部や都市近郊を中心としたサービスエリアであるとしているが、農作物の獣害対策などで山間地から民家周辺への移動を検知するシステムとして適用する可能性は残されている。ただし、端末の電池寿命が1週間あまりと短く、長期の追跡には向かないことを示している。この電池寿命については、現在市販されている携帯電話端末においては1週間程度のずれは出るかもしれないが、大きくは変わらないと考えられる。

これまでに述べたように、中川 [2] や Tsuji ら [3] の研究により、猿は季節の変化に合わせて食物を採っていることがわかった。

また、江成ら [5] や吉田ら [6] の研究により、猿は農作地を食料の供給源として利用しており、人との遭遇を避けながらも農作地の利用を継続している。農作地の利用頻度についても、山に存在する食物の量に依存しており、冬のように山に食物が少ない季節は農作地への依存度が高くなる傾向がある。辻 [7] の研究においては猿の食性の季節的な変動だけでなく、高緯度・低緯度地域などの空間的変異や、群れにおける個体間変異、また、群れ内の社会的立場においても食性が変化するとしており、猿の群れを長期的に観測する必要性がある。

また、大井ら [8][9] や永田 [13] が報告しているように、他の野生動物であるクマやシカとも共通の性質があり、人との遭遇を避けていることや、行動圏を持ち移動しながら食物を得るための行動をしていることから、移動要因や環境要因は野生動物の行動予測にとっては重要なファクターであることがわかる。

行動推定を行うにあたっては何らかの手法による行動のモデル化が必要であるが、これまでに野生動物の行動推定を行うという観点からモデル化を行った例は少ない。江成ら [12] が行っている、猿による農作物被害のリスクマップを作成する試みは希少な例のひとつである。猿の群れの分布や山間部・平地などの地形的な特徴を加味することで農作物被害のリスクマップを作成しているが、市町村単位の大きさであらわす区分になっており、集落単位の大きさとはなっておらず、猿の出現を予測するには範囲が広すぎる。また、時間軸での出現・非出現の議論を行っておらず、リスクの議論にとどまっている。山端 [1] や中村ら [18]、山田 [19] が指摘しているように、猿の農作物被害への対策をする場合には、猿が集落にあらわれた際にはすぐに追い払いを行うことが重要であり、出現する時間に言及しないリスクの提言だけでは対策効果が低いと考えられる。

猿の行動のモデル化では、江成 [15] によって行われた猿の群れの拡大・非拡大を予測する研究があるが、群れの拡縮が農作地への出現を直接的にあらわすわけではなく、人間との関わりの中で猿の行動圏が変化していくことを山端 [11] が論じている。ただし、江成による研究は行動のモデル化に住宅地の割合、標高、植生、平均気温などの環境要因を用いている点で本論とも関連があるといえる。

また、猿以外の動物の行動をモデル化した例として、有本ら [16] によるクマの行動の研究がある。有本らの研究ではクマに GPS 首輪を取り付け、行動を記録した後に、行動についての区分を推定する研究であり、クマの移動先を予測するような行動モデルの研究ではない。

追い払いに関しては自動化に関する議論もあるが、田中ら [21] が実験で示しているように猿は学

習能力が高いために自動的な追い払い手法ではすぐに馴化してしまい、効果がなくなるとしている。

モデル化などの手法に頼らずに、猿の行動を把握する手法として宇野ら [22] や佐伯ら [23] が述べるように GPS や VHF の電波を利用した方法があるが、GPS は行動を記録したのちに装置を回収するために即時性がなく、VHF は電波強度を利用するために即時性はあるものの、位置の検出精度が不安定である。また、装置の電池寿命の問題から交換の手間が必要であり、動物に取り付けるにしても捕獲が難しく、群れの規模に対して必要な頭数が捕獲できるかどうかの問題もある。もし群れの規模に対して電波の発信機を取り付け数が少ない場合、群れの行動を見逃してしまうことにつながり、被害対策効果は減少すると考えられる。

以上のように、猿は農作地を餌場として利用しはじめると、季節的な利用頻度の変化はあるものの、農作地を含んだ行動圏を形成してしまう。農作物の被害対策には住民による追い払いが最も効果的であるが、猿の行動を即時的に知ることは難しく、また、移動などの予測に対するモデル化もほとんど行われていない。そこで本論では、最終的には発信機などの機器を必要としない猿の行動のモデル化を行うため、猿の移動についての調査を行い、移動要因や環境要因をパラメータとして機械学習による猿の出現予測を行う。

2.2 マルコフモデル

本節では、状態変化のモデル化に広く利用されているマルコフモデルの先行研究について述べる。

現在の状態が直前の状態によって決まる性質を持つ事象をマルコフ過程といい、マルコフ過程を確率的に表現したモデルをマルコフモデルという。マルコフモデルには、取り扱う状態が離散的であるマルコフチェーンや、状態遷移を直接観測できない内部状態を表現する隠れマルコフモデルなどがある。マルコフチェーンは状態遷移について十分な回数の観測を行うことで、ある状態から別の状態への遷移確率を単純な乗除で計算することができる。

本間ら [25] はマルコフモデルの 1 種である隠れマルコフモデルを用い、行動パターンを分析する対象としてイベント会場を訪れた人物の動線を選び、行動パターンを分類・認識する手法を提案している。報告によると、近年は防犯カメラや各種の施設におけるウェブカメラなどの設置が進んでおり、このカメラ映像を利用したサービスが多彩に検討されている。その中でも、カメラ映像における人物の動線を用いて、個人の移動の移り変わりが推定できれば、展示会場などで来訪者が興味を引くブースなどの情報を主催者が把握でき、質の良いサービスを提供できる可能性があるとしている。研究では、データセットとしてイベント会場を訪れた来場者の動線を映像からデータ化し、マルコフモデルの一種である隠れマルコフモデルを用いて動線データのモデル化を行っている。隠れマルコフモデルは時間軸方向の変異と伸縮に対してロバスト性を有しており、人間の行動パターンのモデル化によく用いられている。隠れマルコフモデルは、複数の状態とそれら相互の状態遷移確率をあらかず単純マルコフモデルに、各状態からのシンボルの出力確率を加えたモデルである。マルコフモデルには、大きくわけてエルゴード型マルコフモデルと left-to-right 型マルコフモデルの 2 種類が存在する。エルゴード型マルコフモデルとは、どの状態からでも別の状態に遷移できるマルコフモデルのことであり、こういった性質をエルゴード性という。left-to-right 型マルコフモデルとは、ある状態から別の状態に遷移してしまうと、前の状態には遷移できなくなるマルコフモデルのことである。実験では単一のモデルから複数のパターンを学習できるとしてエルゴード型マルコフモデルを使用している。実験環境として研究室のオープンキャンパスをイベント会場に見立て、来場者の動線データを取得している。動線データの取得結果については、隠れマルコフモデルは学習データが十分に確保できない場合は確率分布の近似精度が低下するために特徴的なパターンの学習とモデル化が困難になるという Zhang ら [26] の知見を取り上げ、来場者の少ないブースにおける動線データを省いている。また、隠れマルコフモデルは状態数や状態遷移によって定義される構造に強く依存する [27] ために、本間らはイベント会場を格子状に区切った領域に隠れマルコフモデルの各状態を割り当て、状態数を変化させたとき

の認識率を調べている。隠れマルコフモデルは、他に青木ら [28] が個人の行動パターンを認識する手法として用いており、報告では日常的な行動と非日常的な行動を検出することができたので、行動の例外パターンの認識に有効であることを述べている。

川口ら [29] は left-to-right 型マルコフモデルの適用例として、水族館内における人間行動を分析している。報告によると、集客施設における利用者の行動と施設内で提供している情報をリンクさせて分析した研究事例は少ない。そこで水族館において情報端末機器からのデータ入力と被験者の見た水槽位置・時刻・参照した情報等を組み合わせ、被験者がどのようなルートで館内を移動していったかを遷移確率で示している。水族館のように、ある水槽から別の水槽に移動するような周遊ルートが決まっている場合、どの水槽同士でも自由に行き来できることはほとんどないと考えられ、マルコフモデルとしてエルゴード型ではなく left-to-right 型を採用し、周遊コースごとの遷移確率を行列形式で示している。また、斎藤ら [30] も商業地において複数ブロックの歩行者の流入・流出人口を調査し、歩行者の回遊行動を隠れマルコフモデルを用いて推定しようとしている。

浅原ら [31] も歩行者の行き先を推定する研究を行っている。研究では、歩行者の位置データを蓄積することの研究意義として、例えば作業員の位置データにはその作業の特性、顧客の位置データにはその店舗における顧客の嗜好、個人の位置データには個人の特徴が蓄積され続けるため、蓄積されたデータから新たな情報を抽出していくことを説いている。データセットとして展示会場をシミュレートしたデータを用いて、まず歩行者の動線をクラスタリングし、この結果を状態として遷移確率の計算を行っている。既存の遷移状態予測の手法である隠れマルコフモデルの問題を以下のように説明している。隠れマルコフモデルにおいて遷移するのは空間状態ではなく内部状態であり、空間状態は内部状態により出力されるシンボルである。そのため、歩行者の空間状態は直前の空間状態に依存しないことになる。しかし、実際には歩行者は隣接する空間にしか移動できないため、直前の空間状態は歩行者の行動に大きく影響を与えるはずである。そこで研究手法では、歩行者の嗜好などの1つの内部状態が決まれば、それ以降の歩行者の行動は直前の空間状態によって決まるという混合マルコフモデルを提案している。混合マルコフモデルの特徴は複数のマルコフモデルを重ねたモデルであり、空間的な制約を表現するのに適しているとしている。隠れマルコフモデルは時間変化を表現するのに適したモデルであり、混合マルコフモデルは内部状態がほとんど変化しない場合や個人の嗜好など時間によらない特徴に依存する場合に適したモデルであるとしている。しかし、混合マルコフモデルのデメリットとしては、内部状態が変化しないと仮定したモデルであるため、内部状態が変化するほどの長期間の状態遷移には適しないと述べている。実験では空間の状態遷移のみを計算に用いた単純マ

マルコフモデル，内部状態が多数存在するとした隠れマルコフモデル，内部状態が1つだけとする混合マルコフモデルでの比較を行っている．実験の結果として，隠れマルコフモデルがもっとも性能が悪く，単純マルコフモデル，混合マルコフモデルの順に性能が良くなっていることを示している．隠れマルコフモデルは前述の通りに空間状態同士の依存関係がないため，空間状態の遷移には適さないことを実験から明らかにしている．また，浅原ら [32] は混合マルコフモデルの研究をさらに進め，内部状態を複数仮定するが内部状態は自由に遷移するのではなく，歩行者の嗜好などによりグルーピングされた内部状態間でのみ遷移できるモデルとして混合自己回帰マルコフモデルというモデルを提案している．

マルコフモデルの状態に滞在時間の特性をもたせたマルコフ移動モデルについて岡本ら [33] は丁寧に解説し，シミュレーションによる解析を行っている．報告によると，移動体通信は1つの閉ざされたサービスエリアからなり，サービスエリアは複数のゾーンによって構成される．移動機はゾーンの中心の代表点に，ある分布に従う時間だけ滞在すると仮定し，ゾーンやサービスエリアは簡便のために形状が四角形であるとする．移動機はゾーンの滞在時間が終了した際には瞬時に隣接ゾーンへ移動するとし，どの隣接ゾーンを選択するかの確率（遷移確率）はあらかじめ決まっておらず独立とする．移動機の滞在ゾーンは移動機にとっての状態としてみる事ができるとしている．また，移動機は通信状態を内部状態として持ち，滞在とは別に独立した時間分布にしたがって通信中状態と非通信中状態を交互に繰り返すとしている．

石井ら [34] は移動体通信網の移動機のゾーン間の移動について単純マルコフチェーンモデルを適用している．報告によると，このマルコフ移動モデルは移動体通信網のサービスエリア中の隣接ゾーン間を各移動機が独立な二次元ランダムウォークを行うが，ゾーンにおける移動機の滞在時間の分布関数と隣接ゾーンへの遷移確率をパラメータとして移動機の移動が特徴づけられるとしている．研究では移動機の状態パラメータを計算方法で示し，複数のゾーンを1つのゾーンとして併合できるかどうかの議論を行っている．研究のまとめとして，閉じた複数ゾーンにおいて隣接する外部ゾーンへの遷移確率が初期ゾーン，最終ゾーンに関わらず一定であれば，その閉じた複数ゾーンは併合可能であるとしている．

次に石井ら [35] は先に説明したマルコフモデル [34] を2重マルコフチェーンに拡張し，移動機の移動を模擬する試みを行っている．先のマルコフ移動モデルではゾーン間の移動が単純なランダムウォークであり，道路に沿った移動等の表現が難しいことを問題としてとりあげており，2重マルコフチェーンモデルを適用することで過去のゾーン滞在履歴を考慮した特徴を持つマルコフモデルを生

成できるとしている。このように2重マルコフチェーンに拡張する利点を以下のように述べている。単純マルコフチェーンモデルでは、ほぼ同一地点で互いに逆方向に移動する場合にそれぞれの移動に別のゾーンを設定してモデル化を行う必要があり、ゾーン数が増えてしまう。これに対して2重マルコフチェーンモデルを用いたゾーンのモデル化では、同一ゾーンにおいて逆方向の移動を表現することが可能となる。また、交差点のように互いに逆行する移動機が同一方向に移動するとき、現在位置から同一の確率で移動することは、単純マルコフチェーンモデルではあわせないとしている。交差点では直進、左折、右折でそれぞれ遷移確率が違うため、ひとつのモデルでは進入方向が逆になったときに左右が反転してしまうが、2重マルコフチェーンモデルならば十分に表現できるとしている。

マルコフモデルは逐次的な意思決定モデルではあるが、なぜそのモデルの状態同士の連続的な関連性が得られたのかということを示明的に説明できるものではない。兼田ら [36] は多用事・多立ち寄りの特徴とする空間行動モデルを回遊行動モデルと表現している。人がどこかの地区を訪れた後に、その地区内での移動先を自由に決める場合には、その時点の選択が前の選択履歴に依存しないという意味でマルコフ性を有するとしている。しかし、回遊行動にはスケジューリングの概念が存在するため、既存のマルコフ型回遊行動モデルだけでは説明しきれない行動もあることに言及している。西野ら [37] は観光周遊行動を対象として、複数目的地の組み合わせ決定に関する分析を行っている。分析では、京都市の観光地を自動車でもあったか公共交通機関でもあったか、はじめての観光か2回目以降の観光かの2×2の分類で整理している。研究によると、自動車か公共交通機関かという周遊手段に関しては、移動の自由度に関する問題であり、路線や駅などの制約を受けない自動車利用の場合は複数の観光地に分散して訪れるとしている。また、2回目以降の観光の場合、一日に訪れる観光地の数が少なく、観光地の選択に多様性があるとしている。つまり、個人の嗜好が反映され、気に入った箇所のみ訪れるようになっているとも考えられる。

笠原ら [38] は単純マルコフモデルに環境要因を考慮した観光行動モデルの構築を行っている。報告によると、観光地を周遊する場合には、一般的に旅行者は自分の嗜好に合った観光スポットを見つけ、旅行の制限時間内にどのスポットをどのような順番でまわるかという旅行計画を組み立てる。観光行動は、主体となる旅行者がだれか、次に訪問するスポット、旅行者の嗜好、季節・時間帯・天候などの観光地が属する地域の環境要因の4つの要因によって構成されるとしている。観光行動は既存の研究 [39] で単純マルコフモデルなどで記述されているが、すべての遷移において遷移に影響を与える因子の度合いは等しいと仮定している。報告によると、単純マルコフモデルを用いた手法では直前に訪問したスポットのみを遷移影響因子としている場合が多いが、実際には遷移影響因子の度合いは

スポット毎で異なるため、複数の遷移影響因子の組み合わせとしてスポット毎にモデルを記述したほうが精度の高いモデルを構築できるとしている。季節や時刻を因子として取り上げたマルコフモデルの研究は山崎ら [40] によっても行われているが、笠原らの研究は山崎らの研究を基にして発展させたものといえる。実験の環境要因として、季節・時間帯・天気・地域・方角・距離を選んでいる。環境要因の組み合わせを用いて遷移の条件付き確率を計算し、遷移元のスポットごとにどの環境要因の組み合わせがもっとも条件付き確率が高くなるかを調査し、高くなった環境要因の組み合わせを採用している。環境要因の組み合わせによる条件付き確率を計算する際には、環境要因毎に分割数を変化させ、どの分割数が効果的かも調べている。研究では環境要因モデルの性能を評価する実験として、スポット間遷移予測を行い旅行者が次に訪れるスポットを正しく予測できるか精度を比べている。実験では、京都市の観光地を訪れた旅行者を対象としており、leave-one-out 交差検定での予測精度の平均値を結果として示している。leave-one-out 交差検定法とは、データセットから 1 つだけを予測対象、他のデータを教師データとして用いるというデータ選択をデータセット全体で繰り返す検定法である。実験によると、単純マルコフでの予測結果に比べて、全スポットで共通の環境要因を用いた予測結果、スポット毎に違う環境要因を用いた予測結果の順に予測精度が上がったことを示している。また、選ばれた環境要因として多かったのは天気・季節・時間帯であり、少なかったのは方角・距離・地域であったとしている。結果についての考察は笠原らは行っていないが、本論が対象としている猿の出現予測においても環境要因として天気・時間帯を用いており、生物である人間の行動特性が環境要因によって変化するというのは非常に興味深い。

また、丹後ら [41] もマルコフモデルの適用において、遷移に与える影響について個人差を考慮する必要性を説いている。報告によると、個人の時系列的な状態遷移を重ね合わせることで集団の時系列的な状態遷移を分析するためにマルコフモデルを用いる場合、従来は全ての個人の遷移確率が等しいという仮定のもとで個人の遷移回数を足し合わせることで集団としての遷移回数を求め、集団としての遷移確率を推定していた。研究では、実際の問題に適用する場合には全ての個人の遷移確率が等しいという仮定は成り立たない場合の方が一般的であるとして、以下の問題を指摘している。

1. 個人差のあるデータを足し合わせるのは、異なる母集団をもつ複数の集団が同一の母集団に属しているというような適用方法になってしまう。
2. 個人に合った予測が必要な場合において、従来の仮定では集団としての遷移確率を 1 つだけ求めるようなものなので個人の性質を遷移確率に反映できない。

小澤ら [42] は生活者の行動をマルコフモデルを使って再現し、モデルから家庭エネルギーの時系列

変化をシミュレートする研究を行っている。報告によると、近年では再生可能エネルギーを利用した小規模エネルギー発電の利用が増えているが、大規模エネルギー発電に比べると電力の平滑化が難しく、需要の予測が必要であるとしている。また、小規模エネルギー発電の需要は個人の活動に大きく依存するため、生活行動をモデル化し、モデルから需要を予測することが有効なアプローチであるとしている。研究では既存の生活行動の予測手法として穴埋め問題による手法を挙げているが、研究手法では行為の順序や連続性に現実感のない生活行動が再現されることを問題として挙げており、マルコフチェーンモデルにより逐次的に生活行動を再現していくことで改善されると述べている。実験ではシミュレーションによる生活行動モデルの再現を行っており、数値的な評価はしていないものの、現実感のない行動の再現はされにくくなっていると評価している。また、4人家族を想定した実験を行っているが、個人の行動スケジュールを独立して予測しているため、たとえば食事の時間がずれるなどの問題もあり、モデルの改良は必要であることを述べている。

高山ら [43] は吸収マルコフチェーンモデルを用いて交通量の推定を行っている。吸収マルコフチェーンとは、ある状態から別の状態への遷移ができなくなる状態、つまり、他の状態への遷移確率がゼロになるようなマルコフチェーンのことをいう。実験では4つの交差点が格子状に配置されている道路を対象として、それぞれの道路の端と、交差点に挟まれた道路中に発生源と吸収源をもつ街路モデルを設計し、交通量の推定を行っている。

奥村 [44] は、マルコフチェーンモデルの極限遷移確率について、PageRank アルゴリズムとの比較をまじえながら解説している。マルコフチェーンモデルの極限遷移確率とは、吸収過程がなく再帰的で周期を持たないマルコフチェーンモデルにおいて、無限回の遷移を行った場合に状態の遷移確率が収束するときの確率のことをいう。PageRank アルゴリズムは Page ら [45] によって開発された Web ページのスコアリングアルゴリズムである。PageRank アルゴリズムはユーザが訪れる確率が高い Web ページほど良質なページとして高いスコアを与えるが、PageRank アルゴリズムのスコアリングは Web ページをマルコフチェーンモデルでモデル化したときに無数のユーザがページ間の遷移を行ったときにどのページを訪問するかの確率として定義され、マルコフチェーンモデルにおいて極限分布を求めることと同等であるとしている。また、実際の Web ページは他の Web ページへのリンクが張られていない Web ページ（吸収過程）が存在するが、PageRank アルゴリズムは、ある確率でユーザが任意の Web ページにジャンプすることで極限遷移確率を求めることができるとしている。

原口らは [46] センサの見通しを考慮した確率的モデルにより、ロボット用の障害物地図の作成を行っている。報告によると、ロボットが安全に移動するためには周囲の状況を知る必要があり、これ

までに周囲環境を地図として推定するための研究が行われている。地図の推定方法としては大きく分けて2種類あり、障害物の特徴点を検出する方法と、地図上の地点毎の障害物の有無を検出する方法がある。前者の問題点として、センサの観測値が障害物のどの特徴点を観測したものを同定する必要があり、また特徴点の数も限定的であるために適用が難しいとしており、後者はレーザー距離センサや画像センサで広い範囲を探索することができるためにどの場所を観測しているのか同定しやすいとしている。センサを使って障害物を検出する場合、一般的にセンサは観測誤差を含むため、確率的に地図を表現する必要がある。報告によると、従来の方法では以下の2つの大きな仮定を行っている。

1. 地図上のそれぞれのセルに独立した観測値が得られるので他のセルには影響しない。
2. 地図上のそれぞれのセルにおける障害物の存在は独立している。

2つの仮定により、障害物による見通しの隠蔽があった場合にはその後ろに存在する障害物が検出されず、セルにまたがった障害物が存在するときにそれぞれのセルで障害物が存在すると検出されてしまうために障害物の分布が広がってしまう問題があるとしている。また、地図上のある一定領域内のセル全てに障害物が存在する（または存在しない）確率は、すべてのセルの存在確率の積であらわされるため、単位セルのサイズが小さくなればなるほど存在確率が小さくなってしまふ。障害物や空き領域は現実にはある程度の大きさが存在するため、セルのサイズが十分小さければ、隣にセルにまたがって存在するとして、原口らはこれを空間の連続性と呼び、センサの視線上に連続するセルをマルコフチェーンとみなすことで障害物の推定を行っている。

以上のように、人間の移動や行動をマルコフモデルで分析した研究はいくつもある。

本間ら [25] によるイベント会場での動線から展示物間を移動するモデルであったり、川口ら [29] による水族館の展示水槽を移動していくモデルがあるが、マルコフモデルにおいては状態遷移の方向が限られる left-to-right 型であれば、比較的モデル化は容易であると考えられる。もし猿の移動が left-to-right 型であれば、人の行動予測のように高い精度を出すことは可能かもしれないが、猿はどこに出現するかわからず、出現する集落の順番が決まっているわけでもないために、どの状態からでも別の状態に遷移できるエルゴード型とせざるをえない。また、高山ら [43] のように吸収型マルコフモデルを想定できれば予測の精度を上げられる可能性もあるが、猿の群れが特定の地域1ヶ所に留まり続けるという報告はこれまでにされておらず、吸収型を無理に想定するならば、捕獲による頭数削減をあてはめざるを得ないが、捕獲による群れの規模変更はモデル化の複雑さを増す可能性が高く、予測精度の向上につながるとは考えにくい。

浅原ら [31] が述べているように、マルコフモデルの状態遷移を引き起こす内部状態とは、人の行

動でいえば人の嗜好であったりそのときの目的である。内部状態が決まれば状態遷移をモデル化し、行動予測の精度を上げることが可能である。兼田ら [36] や西野ら [37] が報告しているように人の行動は意図を解釈しやすく、モデル化しやすいともいえる。

マルコフモデルの一種である隠れマルコフモデルは内部状態の遷移が表層にあらわれる状態を決めているとするモデルであり、本間ら [25] や斎藤ら [30] は隠れマルコフモデルで人の移動をモデル化しようとしている。しかし、浅原ら [31] が述べているように、隠れマルコフモデルは内部状態の遷移だけが依存関係にあるため、表層の状態間には関係性がないと仮定するため、実際に場所を移動していくような行動では空間的な移動の制約が存在するため、内部状態の遷移だけでは表現しきれないとして、移動のモデル化には隠れマルコフモデルよりもマルコフチェーンモデルの方が向くことを実験で示している。

マルコフチェーンモデルを移動の状態遷移に使った研究として、石井ら [34][35] による研究があるが、単純マルコフチェーンモデルでは移動の方向性を表現しきれないが、2重マルコフチェーンモデルならば方向性の表現まで可能であるとしている。

西野ら [37] は人間の観光周遊行動を対象として、複数目的地の組み合わせ決定に関する分析を行っているが、交通手段や何回目の観光かによって訪れる観光地の傾向が変わることを述べている。西野らの報告は人の行動に対する知見であり、本論が対象とする猿の行動にそのまま当てはめることは容易ではないが、猿の行動を示唆する材料にはなり得る。ここで、交通手段の違いは移動の自由度に相当し、移動の自由度とは猿の群れの行動スピードであると想定できる。行動スピードは群れの規模やコドモ・アカゴ猿の割合によって変化する可能性が考えられる。猿の行動に関する先行研究はすでに述べたが、群れの行動スピードに関する調査は少ない。また、嗜好が反映され気に入った箇所のみ訪れるようになるというのは、猿にとっての食物の地理的分布状況により、特定の場所によく出現するようになる可能性があり、こちらに関しては先行研究に関する報告で述べた [4][6][7]。

すでに説明したように、人間の行動予測に使われるマルコフモデルは逐次的な意思決定モデルではあるが、なぜそのモデルの状態同士の連続的な関連性が得られたのかということをも明示的に説明できるものではない。笠原ら [38] は人間の観光行動をマルコフモデルを使ってあらわそうとしている。Ashbrook ら [39] の研究では単純マルコフモデルにおいて遷移に影響を与える因子の度合いは等しいと仮定しており、直前に訪問したスポットのみによって状態遷移が決まるとしている。しかし、笠原らはこれを否定し、季節・時間帯・天候などの環境要因が移動に影響を与えることを調査しており、本論の移動要因・環境要因によって猿の行動予測を行えるという主張とも合致する。

丹後ら [41] はマルコフモデルにおける状態遷移には個人差が存在する事実を取り上げ、個人の観察結果が集団の状態遷移をあらわすわけではないことを説いている。また、小澤ら [42] は家族における生活行動モデルをマルコフモデルを使ってあらわそうとしているが、食事など家族がそろって行動していると想定できる場面においても個人の特徴が優先されてしまうと述べているため、集団のモデル化の難しさがわかる。猿は群れ単位で行動することが多いため群れをひとつの単位として取り上げるが、猿の群れひとつひとつが規模や分布に応じて特徴を持つとすれば、猿の移動に関するモデル化が進んでいない現状では、群れについて基礎的な調査から始める必要がある。

以上にまとめたように、人間の移動に関する行動モデルはマルコフモデルを使って表現されることが多い。人間の行動はデータのサンプリングが比較的容易であるために、十分な量のデータでモデル化を行うことができる。また、観光地やイベント会場などを対象にしているものが多く、人間がどういった意図で行動しているかということを想定しやすい。猿の行動の場合、基本的には餌の採集のために移動するということはわかっても、他の群れとの住み分け、山に存在する餌の量、集落での敵（人間）との遭遇など、観測しきれない事象が多く存在するため、猿の行動の意図を想定することは難しい。そこで、猿の行動を予測するためには基礎的な研究からモデル化を試みる必要があり、本論では、マルコフチェーンモデルを採用する。

2.3 ベイズ推定

本節では、確率的な予測モデルを構築可能な手法である、ベイズ推定の先行研究について述べる。

ベイズ推定はベイズの定理を用いて、ある事象に基づきその原因となった事象を推定する手法である。事象 A が発生する確率を事前確率、ある事象 X の基で事象 A が発生する確率を条件付き確率（事後確率）とするとき、事象 A の基で事象 X が発生する確率を観測しておくことで、条件付き確率を求めることができる推定方法である。

小林ら [47] は自律搬送システムの戦略決定アルゴリズムにベイズ推定も用いている。報告によると、ベイズ推定の利点として、既知な情報を事前分布とし、獲得した環境情報から次の環境情報である事後分布を推定することができ、環境情報を無駄なく使用することで素早く次の環境に適応することができることを挙げている。実験では環境情報を特徴として搬送システムの各エージェントが観測した局所的な情報を交換することによって、全体の環境情報を構築している。全体の環境情報と各エージェントが持つ利得から、最も期待値が高いと考えられる戦略を実行するアルゴリズムを作成している。また、環境から取得したデータが現在持っている時系列データから生起する確率が非常に低い場合は、環境が変化したと認識して過去の環境データを破棄し、新たな環境情報を構築する戦略をとっている。これもベイズ推定の結果を用い、事前分布と事後分布の整合性をみることで環境の変化を検知している。

大園ら [48] は買物交通手段選択モデルの構築に最尤推定とベイズ推定を用いて両者を比較している。実験によると、交通手段選択モデルの推定結果では最尤推定とベイズ推定においてパラメータの符号や比率について類似性が存在している。そのため、どちらの推定法を用いても類似した結果が得られる可能性が高いと述べているが、最尤推定に比べてベイズ推定はパラメータ推定にかかる時間を 1/10 程度に短縮できるほか、最尤推定は尤度関数の最大化において局所解が発生するなど困難な問題に直面する場合があるとしている。また、ベイズ推定におけるパラメータの分散を取り上げ、個人間の異質性を検討できるとしている。もしパラメータの分散が大きい場合、そのパラメータは個人間で大きく異なっており、この情報を計画情報として用いる意義を述べている。

Ito ら [49] は WiFi を用いた位置推定の方法である FingerPrint 方式をベイズ推定を用いて行っている。FingerPrint 方式とは、事前に Received signal strength indication（以下は RSSI とする）をデータベースに記録しておき、位置を推定したい RSSI に対してデータベースを参照することでどこで発生した RSSI かを調べる手法である。Ito らは電波を利用した既存の位置推定法の手法の問題として以下のように多岐に渡って述べている。

1. 位置推定のシステムはユーザの位置のプライバシーについての問題を抱えており、特に屋内の位置推定をする場合、GPS や携帯電話などの屋外の位置推定システムに比べてユーザにとっては非常に正確な位置を把握されてしまうことになり、ユーザ位置が信頼のおける管理者によって運用されるかどうかはわからない。
2. 位置推定のための新規デバイスを導入することは困難を極める。
3. 電波はデバイスの向きに非常にセンシティブであり、容易に RSSI が変動してしまう。
4. 特に室内においては、マルチパスの問題が起き、反射や干渉による RSSI の変動が大きくなる。

マルチパスをはじめとした RSSI の変動を原因毎に切り分けていくのではなく、すべての変動を含めた観測を行い、ベイズ推定を適用することで位置推定を行っている。

伊藤ら [50] はさらに研究を進め、アクセスポイント（受信点）の選択性に関する報告を行っている。報告では、マルチパスなどによる RSSI の変動幅が、アクセスポイントに近い方が大きいことを実験的・理論的に確かめている。結果を受け、一定時間観測したときの RSSI の平均値がしきい値以内にあるアクセスポイントを採用することで位置推定の精度を向上させている。実験では、RSSI の平均値によってアクセスポイントを選択するが、アクセスポイントに近いほうが平均値の変動が大きいため、なにもケアしない場合は遠いアクセスポイントのほうが選択されやすくなってしまふ。遠いアクセスポイントが選択されてしまうことの弊害がどの程度になるかを位置推定の精度から確かめており、平均値の大小で選択するアクセスポイントを変え、平均値が高いアクセスポイントを選択した場合の方が位置推定の精度が高いことを示している。また、RSSI の平均値を計算するための観測時間を変え、観測時間が長い方が位置推定の精度が高いことも示している。理由として、RSSI は同一の位置においてもごく短時間で変動してしまうためとしている。実験では、時間による RSSI の変動の大きさを 1 日にわたって調べており、これによると、人が活動していると思われる 10~22 時には変動幅が大きいため、RSSI の事前観測と位置推定を行う際の観測をする時間がずれてしまうと位置推定の精度に影響が出てしまう可能性を指摘している。

松月ら [51] は屋内での FingerPrint 方式においてベイズ更新をし続けることで最終的には位置推定の誤差を 1m 未満まで抑えている。

高橋ら [52] も同様に FingerPrint 方式をベイズ推定を用いて行っているが、FingerPrint 方式の欠点のひとつを解消する試みを行っている。報告では、無線 LAN のアクセスポイントが一定の出力で電波を送信していることを前提にしていることを従来の FingerPrint 方式の問題として挙げており、最

近のアクセスポイントが省電力モードにより出力を不規則な期間で抑えていることを解決できていないとしている。そこで複数の期間に渡ってアクセスポイントの RSSI を記録しつづけ、この記録を元にベイズ推定で位置推定を行っている。実験によると、特定の日時に記録したデータを評価に使った場合と、全ての期間からランダムに抽出したデータを評価に使った場合の推定精度の比較を行っており、前者では 8 割に満たず、後者では 9 割以上の精度が得られたとしている。実験結果は、特定の日時のデータのみを評価用とした場合には省電力モードなどの影響を受け、モデル用のデータと特性が大きく乖離してしまう可能性によるものであると述べている。そのため、ベイズ推定のような確率モデルでは、モデル作成用の事前調査をできるだけ多く行うことが対処法であることを考察している。

加治ら [53] は FingerPrint 方式のように環境地図を用いた位置推定の問題点を以下のように挙げている。

1. 構築済みの地図からの逸脱や復帰の検出
2. 逸脱を検出した際の環境地図の拡張方法
3. 構築した環境地図内での位置推定精度

上記の問題点のうち、逸脱に関してはアクセスポイントの位置を既知として、どのアクセスポイントの信号を受信したかを検出すればよいとしている。また、電波を利用した位置推定の方法には大きく分けて以下の 3 つが存在し、複雑な電波伝搬環境に対応しやすい学習方式として FingerPrint 方式が使われていることを述べている。

1. 3 点測量方式：基準点となるアクセスポイントの距離を RSSI から計算し、位置座標を求める。
2. 近接方式：電波が受信可能であるかどうかの判断のみで位置を求めるため、位置の推定精度は電波の到達範囲と同じになってしまう。
3. 学習方式：各座標における RSSI を記録しておくことで位置座標を求める。

また、佐藤ら [54] は 3 点測量方式の精度を上げるために、指向性アンテナを使って受信方向に指向性を持たせ、受信方向に存在するアクセスポイントを選択するための方法としてベイズ推定を用いている。

本村 [55] は、ベイズ推定により日常生活行動のモデリングをする研究を行っている。報告によると、行動のモデル化を行う場合には、行動自体の収集だけでなくその原因となった変数を含めた観測データを収集し、得られた大量の変数間の関係から因果関係を導くことが必要であるとしている。研

究では、子供の日常生活行動をモデル化するために、子供に超音波センサをとりつけ、室内で過ごしたときの構造物との相対距離や子供の動く速度などを説明変数とし、行動の正解ラベルは映像から手動で行っている。実験ではベイズ推定を適用するにあたり、確率変数をノードとするグラフ構造であるベイジアンネットを適用することで、推定精度を上げている。結果として、単純な最尤推定では50%未満の推定精度であったものが、ベイジアンネットを適用することで60~80%に上げることができたとしている。

手塚ら [56] はベイジアンネットを車両変更時の運転行動推定モデルに適用する研究を行っている。研究では、車線変更などの運転行動を想定した研究のひとつとして隠れマルコフモデルが使われている例をとりあげているが、運転行動の連続推定を行った場合に操舵開始直後において誤推定が起こる問題点を挙げ、運転行動の前後関係を考慮できるモデルとしてベイジアンネットを適用して連続推定を行っている。手塚らは隠れマルコフモデルに対するベイジアンネットの優位な点として2点を挙げている。

1. モデルパラメータの値自身も確率的に変動するため未知入力に対する過学習を緩和することができる。
2. 連続でなめらかな特徴量系列を生成できるため対象に近いモデルを提供できる。

実験としては、ドライビングシミュレータを用い、車線維持、通常右車線変更、緊急右車線変更の3状態の推定を行っている。入力パラメータとしては、操舵角度の時間変化を与えている。実験によると、ベイジアンネットを用いた提案手法は隠れマルコフモデルを用いた手法に比べて安定して推定できているとしている。また、教師データとテストデータの特徴量の近さをノルムで計算しており、ノルムが小さい、つまり両方のデータの特徴量が似ているほど、推定の精度が高いことも結果で示している。

周ら [57] は移動ロボットの位置決めのためのセンサプランニングのためにベイジアンネットを適用している。ロボットは自身の位置決めをするためにセンシングを行うが、その際にロボットが装備するセンサのエラーやロボットの姿勢の変動により、誤推定が発生するためにユニークな位置決めが難しいという問題がある。そのような不確実性に対応するために確率的なアプローチがとられている。ロボットの位置決めの方法には、初期位置が与えられた状態でセンサ情報によって位置決めを逐次行っていく方法と、初期位置が不明な状態でセンシングを行いながら位置決めを行っていく方法があるが、後者の方が難しいとされる。研究手法としては、初期位置が不明な状態での位置決めを行うためのセンサプランニングの方法にベイジアンネットを用いている。報告によると、ベイジアンネットの

ノード間の因果関係を決めることは、実際の問題では難しい。そこで、一般的にはベイジアンネットの構造設計は人がすることが多いと述べている。ベイジアンネットの構造設計を自動で行う手法を提案し、位置決めの実行時には得られたセンシング情報に基づいて適切なセンシング行動をロボットが自動計画するようにし、シミュレーション実験を行うことで提案手法の有効性を示している。

石田ら [58] はコンピュータの侵入検知システムにベイズ推定を用いている。過去のパターンを予測する手法として状態遷移をモデル化するマルコフモデルがあるが、ネットワークイベントから遷移モデルを作成するのは難しいことからマルコフモデルは使用していないとしている。研究では、攻撃イベントが増加するか減少するかの2クラス問題として捉え、攻撃イベントの周期に着目したモデルと、攻撃イベントの増減の度合いに着目したモデルをアルゴリズムとすることで予測を行っている。また、攻撃イベントの増減の度合いに着目したモデルにおいては、度合いの分割数を変化させたアルゴリズムを用意している。実験では、予測期間を1時間毎としたときにもっとも効果的であったモデルは攻撃イベントの増減の度合いに着目したモデルであり、度合いの分割数は2（つまり、増加するか減少するか）のときであった。考察として、周期に着目した場合では時間帯の異なる世界中から攻撃を受けているので時間の周期性を見出すのが難しく、増減の度合いに着目した場合では度合いの分割数を増やしてしまうとそれぞれの場合における発生頻度が下がるため、分割数が2のときがもっとも精度が良かったとしている。このように、ベイズ推定を適用する場合は、確率的に差異が生じる変数を選び、適切な分割数でモデル化を行う必要がある。

植野 [59] はベイズモデルの事後確率の更新を利用し、eラーニングにおける所要時間データの異常値の検出を行っている。報告によると、異常値の検出の手法には検定に基づく手法や教師ありニューラルネットワークを用いる方法があるが、これらの手法を所要時間データの異常値の検出に適用する場合には、次のような問題があるとしている。

1. 最初から異常なプロセスをとっていた場合、その後は異常でないプロセスが異常であると扱われてしまう。
2. 従来の方法では、時系列データの発生源が同じであると仮定しなければならないが、eラーニングではコンテンツが全て異なるため難易度によって学習時間が変化してしまう可能性があるとしている。

上記の問題に対して既知の所要時間データをもとに新たなデータの出現する確率分布をベイズ予測分布で導き、新たなデータの異質性を検定する手法を提案している。提案手法による利点としては以下の2つを挙げている。

1. 判断材料となるデータが少ない場合は異常値の基準を緩め、データが増えるに従って厳密な基準となっていく特性を持つ。
2. 過去の学習データから各コンテンツの所要時間の平均と標準偏差を逐次計算し、各コンテンツの学習時間を標準化したデータ系列を用いた異常値検出が可能である。

報告では、学習者の理解の有無と所要時間の異常の有無をアンケートでとっており、学習者が所要時間に異常なしと答えている場合であっても提案手法では異常値を検出した場合は学習者が理解なしとしている場合が多いことを報告しており、提案手法による異常値検出が有効であることを示している。

田中ら [60] は、ベイズ推定における特徴量選択において新規の試みを行っている。一般的にベイズ推定は、全てのクラスにおいて共通となる特徴量を選択して識別器を作成するが、報告では特徴選択のための潜在変数を導入することにより、識別器を作成する過程において特徴量が各クラスに有効かどうかを自動的に計算し、各クラスに識別に有効な特徴量の集合を用いて識別器を作成する手法を提案している。提案手法の考え方は、特徴量が独立であるとする、識別に有効な特徴量とはその特徴量の分布がクラスにおいて固有となる特徴であるとする。そのため、特徴量がクラスに固有の分布になるときはその特徴量をそのクラスの識別器に用い、クラスに固有の分布にならない場合はそのクラスの識別器には用いないとしている。また、全ての特徴がクラスの識別に有効である場合は通常のナイーブベイズと同等の識別器になるとして、提案手法の比較にはナイーブベイズを用いている。また、比較手法にはロジスティック回帰や線形 SVM も取り入れている。実験結果としては、田中らの提案手法がどの手法よりも精度が高かったとしている。理由として、従来のベイズ推定ではある特徴量の分布が似通ったクラス間で誤識別が起りやすいが、提案手法ではそういった特徴量はクラス間で共通の特徴量として扱い、別の特徴量を識別に使うことで推定精度を上げていると説明している。

岩村ら [61] は予測結果の偏りに関する考察を行っている。パターン認識においてベイズ推定を適用する場合、クラス間でサンプル数が等しければベイズ推定が有効に働くが、クラス間でサンプル数に偏りがあると尤度の偏りのために認識精度が低下するとの指摘がある。研究では、尤度の偏りを補正することで認識精度が向上するかどうかの検証を行っている。予測分布の偏りに関する理論式を得ることで、尤度の補正を行っている。実験としては、正規分布に従う人工サンプルを用いた2クラス問題を取り扱っており、教師データ数比を固定して次元数を1, 2, 4, 8, 16, 32と変化させる実験と、次元数と片方のクラスの教師データ数を固定してもう一方のクラスの教師データ数を変化させる実験を行っている。また、人工サンプルのクラス間の大きな違いは分散の大きさである。実験によると、分散の小さなクラスの教師データが少ないほど認識率が低下する傾向が見られ、岩村らの方法に

よる尤度の偏りの補正効果が得られている。逆に、分散の大きなクラスの教師データが少ないほど補正効果が少ない傾向となっている。また、次元数が少ないほど岩村らの提案手法の補正効果が大きいとしている。結果を受け、ベイズ推定を用いたパターン認識は、予測分布の偏りが認識精度を低下させているという結論を得ている。

風間ら [62] は大規模な語彙の文脈類似度の計算にベイズ推定の手法を取り入れている。報告によると、100万種類を超える非常に大規模な語彙の類似度を計算するために計算コストを最小に抑える必要があるとしているが、ベイズ推定の計算は一般的には計算量が大きくなるとしている。そこで、文脈プロファイルが多項分布であり、事前分布がディリクレ分布であり、類似度が *Bhattacharyya* 係数であるときに、期待値の計算に解析解を求めることができ、効率的な計算ができるとしている。実験結果では提案手法をベイズ推定としては最も単純な部類に入るとしながらも、従来手法では24CPUコアで8日かかる計算を1 CPU コアで1時間で終わるとしており、適切な分布を与えられれば、ベイズ推定が効果的な手法であることを示している。

木谷ら [63] は歩行者の安全支援を行うことを目的として、複数車両間で通信を行うことで歩行者の位置推定を行う研究を行っている。研究では、歩行者にビーコンをもたせ、車両には指向性アンテナとGPSをとりつける。指向性アンテナによりビーコン電波の到来方向とRSSIを取得し、歩行者までの距離と角度をベイズ推定で予測している。また、複数車両から得られたデータや、歩行者の時系列的な位置推定データを用いることで精度の改善を行っている。実験はシミュレーションで行っており、実験条件として測定機器の精度誤差の大きさにより3条件を設定している。距離や角度の測定精度が最も高い測定機器を使用したと仮定した場合、単一車両だけの受信結果を用いた場合でも歩行者の位置推定誤差が7m程度に収まっており、複数車両の受信結果を用いた場合は1m程度の誤差まで減少している。もっとも測定精度の悪い測定機器を使用したと仮定した場合は、単一車両だけの結果での歩行者の位置推定誤差は20m程度あるが、複数車両の受信結果を用いた場合には2m程度まで減少している。以上のように複数の受信結果や時系列のデータを用いることで、ベイズ推定においても良好な推定精度が得られるとしている。

豊田ら [64] は色の恒常性を記述する手法についてベイズ推定を用いている。報告によると、人間の視覚系には、照明光の影響から人間が観測する対象物の表面色を分離して知覚する色恒常性という能力がある。色恒常性は網膜の錯体の分光感度、照明光の分光分布、物体表面の分光反射率を用いて数式で記述できるが、一般には解は一意に定まらず不良設定問題になっているとされる。ベイズ推定が視覚の不良設定問題とマッチングするという知見に着目し、観測シーンの統計量から分光反射率の統

計量と照明光の色度を推定する問題としてとらえてベイズ推定を適用している。実験によると、従来の色恒常性についての仮説よりも、ロバストな照明光推定が可能であった。従来の仮説では、多くの色知覚現象を説明し得ないことが問題であったが、提案手法ではより正確に説明できるとしている。ただし、提案手法が必ずしも正しく照明光と反射率の情報を分離できるわけではないことを説明しており、ナトリウムランプなどの特殊な分光分布特性をもつ光源に対しては、提案手法だけでなく人間も正しい推定を行うことは困難であるとしている。すなわち、人間が色の知覚を行う場合でも、自然物体の分光反射率や照明光に基づいた事前知識に基いてベイズ推定的に行っていると主張している。しかし、照明光や反射率について物理的な正解値が存在しない条件においても、提案手法は実験的に計測された色知識を良好に再現可能であったとしてベイズ推定の有用性を説いている。

深澤ら [65] はベイズ推定に基づく空間統計学的手法を用いて、生物の分布データの解析を行っている。報告によると、野外の生物は空間的に集中した分布パターンを持つことが多い。理由として、生育に適した環境の分布などの影響を受けるとしている。空間の類似性は、既知の環境要因だけでは説明できないことが多いとしている。研究では、隣接した地点間の空間の非独立性を表現できる空間的ランダム効果を持つモデルを用いて生物の分布データを推定している。空間的ランダム効果は隣接した地点の欠損データを補完し、なめらかなパラメータを決定できる効果をもつ。実験の説明変数として標高・傾斜角度・集水域面積・スカイラインを選び、常緑高木の一つであるアカギの分布域の推定を行った。比較手法としてはロジスティック回帰モデルを用いている。実験の結果、報告で使われたモデルは分布域の推定を良好に行えていることを示しているが、その理由としては未知の要因によるばらつきを空間的ランダム効果が吸収したためと説明している。また、ロジスティック回帰モデルの推定精度が低かった原因として、ロジスティック回帰モデルは未知の要因がないことを仮定しているため、空間自己相関を無視することによるバイアスの可能性を指摘している。

福山ら [66] は地磁気の脈動の開始時刻の決定手法として最尤推定とベイズ推定を適用している。報告によると、地磁気の脈動のうち $Pi2$ 型地磁気脈動は極めて緩慢な変動を示す初動部分においてただ一点を開始時刻として同定する作業は非常に難しいとしており、既存の手法として使われているウェーブレット解析や帯域フィルタを用いた手法では、フィルタをかけたデータを使用するためデータ自体の特性が変化してしまい必要な精度を満たせなかったり、開始時刻を決定するための条件が主観的であったりするなどの問題があるとしている。そこで実験では最尤推定とベイズ推定を適用した2つの手法にて開始時刻を求めている。最尤推定とベイズ推定において、入力データは加工法が違うために同一の条件の元で優劣を論じることは難しいが、提案手法のそれぞれの手法で得られた地磁気

脈動の開始時刻がほとんど同じになったことから、地磁気脈動には特徴的な初動成分があるとしており、今後の研究発展を継続するとしている。また、最尤推定とベイズ推定では、初動成分を未知のパラメータとして扱えるベイズ推定の方が有効であるとしている。さらに、提案手法は既存の手法である固定フィルタの利用をせずに済むため、データ特性が変質してしまう問題が解決できるとしている。

小林ら [67] はドライバーの学習行動を考慮した経路選択行動モデルについて、ドライバーの主観的期待の更新モデルをベイズ推定を使って記述している。報告によると、ドライバーが経路選択のために使っている走行条件としては、期待走行時間、走行時間の分散、走行時間の確率分布等があるとしている。ドライバーが過去の実績に基づいて経路を選択する場合、過去の時点での走行時間の予測結果と実績値との乖離等を説明変数としてモデル化していると説明しており、ドライバーの学習行動とは、このような経路走行時間の予測モデルを経験的に逐次修正するメカニズムのことである。こういった学習行動はベイズ推定の更新にマッチするモデルであり、小林らはこれをシミュレーションで説明している。

位置推定についてのベイズ推定の応用例として Fox ら [68] によるベイジアンフィルタを紹介する。報告によると、どのようなセンサであっても完全といえる精度を提供できるセンサは存在せず、そのようなセンサを複数組み合わせる場合は不確かさの取り扱いが重要となる。ベイジアンフィルタはどのような種類のセンサであっても不確かさを確率分布であらわすことができるとしている。

以上のように、ベイズ推定は確率で表現できる幅広い分野で使われている。

小林ら [47] による自律搬送システムや石田ら [58] によるコンピュータの侵入検知システムの例が示しているように、ベイズ推定は過去の時系列的な特徴量の変化が一定している場合に向く手法といえる。

また、Ito ら [49] や高橋ら [52] が RSSI の空間・時間的な分布の変化に対してベイズ推定を適用し位置推定を行っているように、確率分布で表現できる特徴量をパラメータに使える場合にはベイズ推定が高い推定精度を示す。

しかし、ベイズ推定は過去に発生した事象を確率的に表現するモデルであるために、次のような問題を含んでいる。

伊藤ら [50] や高橋ら [52] が述べているように、教師データとテストデータの観測結果がずれている場合はベイズ推定の適用は難しい。手塚ら [56] も教師データとテストデータの特徴量の近さをノルムで計算しており、ノルムが小さい、つまり両方のデータの特徴量が似ているほど、推定の精度が高

いことも結果で示している。観測結果がずれてしまうことの対策として、ベイズ推定のような確率モデルでは、モデル作成用の事前調査をできるだけ多く行うことが対処法であることを考察しており、猿の行動予測に適用しようとした場合には、行動調査が長期に渡ってしまうことを示している。また、猿の群れは年ごとに規模や年齢構成などが変化する可能性があり、数年単位で調査を継続しても毎年少しずつ傾向が変わってしまう可能性もある。十分長期に渡って調査を行うことで、傾向の変化も平均化される可能性もあるが、平均化された調査データを使うのであれば猿の農作物被害のリスクマップを作成した江成ら [12] の研究と大差がなくなってしまう、出現する時間に言及できないような予測モデルになる可能性が高い。また、長期に渡る猿の群れの調査は、調査機器の電池寿命や調査対象とする猿の捕獲の難しさもあるため物理的にも困難である。

また、石田ら [58] が示すように、ベイズ推定の尤度を計算する際には、特徴量の分割数によっても予測の精度が変わることを検証する必要がある。分割数が増えるほど、尤度を真値に近づけるためには十分な量の観測データが必要であり、分割数を減らすほど、特徴量の解像度が下がることになるため、事象を正しく表現できないと言える。環境要因の中でも気温のように連続値として与えられる特徴量も存在するが、連続値をわざわざ離散値として扱わざるを得ない点で不利である。連続値を連続値として扱える手法が望ましい。

さらに、田中ら [60] や岩村ら [61] は特徴量の次元数に言及している。報告では特徴量の次元数が増えるほど推定の精度が下がるとしており、精度の向上に寄与する特徴量のみを選択することが重要としている。長期的な観測のもとに特徴量を選択できるのであれば有効な確率モデルを作成することができるのかもしれないが、上述のように猿の群れの長期的な観測は難しい。

猿の群れの挙動は非常に非線形であると考えられ、どれだけ長期間に渡った調査であっても、十分といえる特徴量の組み合わせを観測するのは難しいと考えられる。そこで、本論では一般的な確率モデルであるベイズ推定を、猿の出現予測の難しさを示すための比較手法としてのみ用いる。

2.4 Support Vector Machine

本節では、非線形的な予測モデルを構築可能な手法である、SVMの先行研究について述べる。

SVMは、Cortesら[69]によって考案されたとされている。SVMは2クラス問題のための機械学習器として開発された。基本的なアイデアとしては、非線形な入力ベクトルを高次元の特徴空間にマッピングする。この特徴空間では線形な識別平面を構築することができ、この識別平面の空間的な特徴を用いることで機械学習器に高い汎化性能を持たせることができる。Cortesらによる手法が考案される前までは、教師データが特徴量空間でクラスごとに完全に分離できていなければいけなかったが、Cortesらによる手法によって、教師データが特徴量空間でクラスごとに分離されていなくても識別できるように拡張された。提案手法では、特徴空間における識別平面と入力ベクトルとのマージンを計算し、このマージンが最大となるように識別平面を設定することにより教師データがクラスごとに混在することを容認している。また、報告の中で、多項式カーネルを用いて識別平面を決定し、文字認識についての認識精度を他の手法と比較した結果、線形分類器と比べて高い汎化性能を示した。

伊波ら[70]は危険なシステムコールに着目した異常検知手法としてSVMを適用している。報告によると、不正プログラムの検出にはプログラムの振る舞いに基づくビヘイビア型異常検知システムがよく用いられているが、高い検知率を得るにともなって正常な振る舞いであっても異常とされるFalse Positiveの割合が増加することが問題であるとしている。False Positiveが小さいほど検知精度が高いとしてセキュリティ対策ソフトの評価指標のひとつとなっている。提案手法により検知能力が高いこととFalse Positiveの低減ができることを研究の目的としている。実験では、ルールベースの検知システムによりクリティカルと判断されたシステムコールについて、そのシステムコールに至るまでの過去の時系列システムコールからSVMの入力データを作り、クリティカルと判断されたシステムコールが悪意を持ったプログラムによるものかどうかの判断をSVMで行っている。SVMのカーネルには線形カーネルを使用している。実験の結果として、不正なプログラムによるシステムコールの評価用データセットの98.3%を危険なシステムコールとして識別できたとしている。また、通常のプログラムによるシステムコールの評価用データセットについても危険と判断していたものが96.7%であったものが、SVMを適用することで15.0%まで減少させることができたとしている。以上のように、False Positiveを減らす手法としてSVMは非常に有効であると考えられる。

山崎ら[71]はSVMを多クラス問題に適用するにあたり、認識の確信度に関する研究を行っている。SVMは2クラス問題を分類するための手法として考案されたが、これを多クラス問題に拡張したone-vs-all SVMがある。one-vs-all SVMでは識別したい特徴量ベクトルが複数あるクラスに所属す

るかかどうかという SVM を学習しておき、入力ベクトルに対して複数の SVM を個別に適用させ、得られた複数の結果のうち、最も識別平面からの距離が大きい（識別関数値が大きい）クラスに所属すると判別することで多クラス問題を取り扱っている。報告によると、識別関数値から認識精度を推定する研究が行われているとされる。機械学習により得られた認識結果がどの程度の自信をもって出力されたものであるかを示すパラメータを確信度と定義している。Maeda ら [72] の先行研究によると、確信度として候補第 1 位と候補第 2 位の識別関数値の差に注目し、確信度の精度としてはこの差を取る手法が最も良いということを示している。これに対して山崎らは、候補第 2 位から第 n 位までの平均を算出して第 1 位との差分を取る手法と、候補第 1 位から第 n 位までの識別関数値を順に並べたものを新たな特徴ベクトルとして 2 クラス問題の SVM とし、その 2 クラス問題の SVM から得られる識別関数値を確信度とする手法の 2 つを提案している。山崎らは実験として物体認識に用いられる画像セットの確信度を算出しており、実験結果によると、山崎らの 2 つの提案手法は既存の手法に比べて良好な結果が得られたとしている。また、2 つ目の手法は SVM の出力結果に対してさらに SVM を適用することから、データセットが十分に大きくなかったために過学習を起し、最善とはいえない結果となったとしているが、十分な教師データが得られる場合はこの問題が解消される可能性があることを述べている。

小林ら [73] は RBF カーネルを用いた SVM を使用するにあたり、パラメータの最適化を強化学習によって行う研究をしている。SVM には、識別平面を決める際に誤判別を許容するためのコストパラメータと、カーネル関数の形を決めるカーネルパラメータがある。コストパラメータを小さくした場合は誤判別の許容量が大きくなり、コストパラメータを大きくすると誤判別を許容しなくなる。また、カーネルパラメータを小さくすると識別平面が単純な形になり、カーネルパラメータを大きくすると識別平面が複雑な形になる。コストパラメータとカーネルパラメータは大小どちらが良いかという議論は難しく、最もふさわしい解答としては、汎化能力が最大となるパラメータに設定するというものである。SVM のパラメータの決定方法としては、交差検証法や最急降下法があるが、報告によるとこれらの方法は少ない学習回数で解が収束することやアルゴリズムが単純で済むというメリットがある。しかし、これらの方法はパラメータの探索範囲が狭くなりがちであったり、局所解に陥りやすいという問題があるとしている。そこで報告では、パラメータの探索に強化学習を用いることで、パラメータの値をランダムに探索することで広い範囲を偏りなく探索する手法を提案している。報告では、まずコストパラメータを固定し、RBF カーネルのカーネルパラメータを探索する実験を行っている。実験の目的としては、パラメータがひとつであるので強化学習を用いて探索する必要はないが、

提案手法が実用可能かどうかの検証を行うためとしている。実験によると、勾配法では局所解に陥りやすい箇所であっても、強化学習を用いた方法では広く探索することで対処可能としている。次の実験として、コストパラメータとカーネルパラメータの両方を探索する実験を行っている。通常のパラメータ探索では、この両方のパラメータを決定する必要があるため、こちらの実験のほうが実用可能かの検証になっているといえる。実験結果によると、はじめは大半の識別率が30%程度であったが、識別率が収束する頃には90%程度が得られたとしている。以上のように、SVMにおいてはパラメータの設定が識別率に非常に大きく寄与する。

青木ら [74] は SVM のための特徴量の選択手法についての研究を行っている。Vapnik [75] により SVM は特徴次元数を削減しなくても過学習が避けられるとされているが、Guyon ら [76] による研究では特徴選択を行うことで SVM の識別精度を向上させることができるという報告がある。特徴選択を行う方法として Whitney [77] による特徴量を最小限から徐々に増やしていく Sequential Forward Selection と、Marill ら [78] による特徴量を最大限から徐々に減らしていく Sequential Backward Selection があるが、青木らによると、どちらの方法も特徴量の組み合わせを全て試せるわけではないため、最適解が得られる保証はないとしている。そこで青木らは山崎ら [71] と同様に SVM の識別関数値（識別平面とサポートベクターとのマージン）に着目し、識別関数値を評価値とする特徴選択の手法を提案している。SVM の識別平面を決定する際に反対側のクラスに入ってしまったサンプルは識別平面のマージンを決めるにあたってのノイズとなってしまう。そこで青木らは反対側のクラスに入ってしまったサンプルに対してペナルティを与えることで信頼性の高いマージンを決定できるとして、提案手法による特徴選択を行っている。実験によると、人工データとして有効な次元2つと一様乱数で構成されたノイズとなる次元数98のサンプルを用意し、提案手法によるマージンを用いて特徴次元を削除していった結果、有効な次元2つのみが残る、かつマージンの評価値が最大になったとしている。実験の対照手法としては通常 SVM のマージンを基準に特徴次元を削除しているが、既存の手法では有効な次元2つのみは残ったものの、マージンの評価値は最大にならなかったとしている。また、実データを用いた実験も行っており、提案手法ではマージンの評価値が最大になるときの次元数と SVM の識別率が最大になる次元数が一致した。既存の手法ではマージンの評価値と SVM の識別率が最大になる次元数は一致せず、既存の手法の方が提案手法よりも SVM の識別率が数%高かったとしているが、提案手法は SVM の学習過程のみで特徴次元を決定することができるが、既存の手法では学習と識別の過程を行わないと特徴次元を決定できないため、計算コストにおいて提案手法の方が1/20程度で済み、実用的であるとしている。

ただし、特徴次元数の削減については有効ではあるとするものの SVM の識別性能の向上効果は小さいとする Melgani[79] の報告もある。機械学習の分野では、特徴次元数が増えすぎると、識別性能が下がる可能性がある。Melgani の報告によると、特徴次元数を 20 から 200 まで変化させたときの結果として、K-最近傍法では分散で 2 程度の識別性能の劣化があったが、RBF カーネルを用いた SVM では分散が 0.3 であったとし、次元数を最適化する影響は小さいとしている。

鈴木ら [80] は人の行動を認識する研究を通して、ラベルなしデータセットのラベルを適応的に識別し、それらのラベルなしデータセットから識別器を学習する手法を提案している。SVM は一般的に教師あり学習と言われる手法の識別器であり、教師あり学習では分類したいクラスそれぞれに正解ラベルが付与されたデータセットが必要である。正解ラベルが既知の状態で作られたデータセットでは、新たにラベルを用意する必要はないが、実世界におけるデータセット、特に測定環境の影響を受けるようなデータセットでは、測定環境ごとに正解ラベルを付与する必要がある。こういった作業は人的・時間的にコストが高くなりやすい。報告によると、Zhou ら [81] が示しているように、ラベルありデータセットが不足する場合には手動でのラベル付けが不要なセルフトレーニングを用いることができるとしている。セルフトレーニングとは、すでに存在するラベルありデータセットをもとにして、ラベルなしデータセットに自動的にラベル付けを行っていく手法のことである。しかし、前述のように測定環境が違うデータセットの場合はこのラベル付けが正しく行われずに識別の精度が下がる可能性があるとしている。そこで、十分な量のあるラベルありデータセットを教師データとして利用して、ラベルなしデータのみからなるデータセット用の識別器を作成するクロスデータセットという方法をとっている。また、クロスデータセットで学習を行う場合、ラベルありデータセットとラベルなしデータセットの特徴分布が異なり、ラベルの予測が正しく行われぬ可能性が高い。報告では、両方のデータセットで認識能力が高い特徴次元のみを用いて識別器を作成することでラベルの予測精度を上げている。提案手法では、まずラベルありデータセットを教師として SVM の識別器を作成し、識別器を用いてラベルなしデータセットの分類を行う。分類を行った際に SVM のライブラリのひとつである LIBSVM[82] が備える確率推定機能をスコアとして、どのクラスに属する確率が高いかというスコアの算出を行い、スコアの高かったデータを新たに教師データに加えるという手順を繰り返すことでラベルなしデータセットにラベルを付与している。以上のように、SVM をはじめとした教師あり学習では、正解ラベルの付いたデータセットをどのように準備するかということが研究の第一歩となる。

LIBSVM は Chung ら [82] によって作成された SVM のライブラリのひとつであり、多項式や RBF

など各種のカーネルを備え、線形問題、非線形問題、1クラス分類による異常値検出や2クラス分類・多クラス分類に対応するなど非常に汎用性の高いライブラリとなっている。

吉谷ら [83] は SVM を用いて新聞記事からのプロフィール情報を抽出する研究を行っている。情報抽出の分野では、人名などの種類に対する単語列を抜き出す固有表現抽出と、人名などの種類を表として表現したときに同一人物など同じものに対する表をまとめる情報統合の研究が盛んに行われている。報告によると、従来の情報統合はすでに表になっている情報しか扱っておらず、通常の電子文書から表形式の情報を作成するようなシステムが必要であるとしている。そこで、テキストから表の項目となる単語を抽出し、それらの項目を対象となる人物ごとに整理する手法に SVM を利用している。報告では抽出精度の評価指標として、Recall, Precision, F 値の3つを用いている。Recall とは、あるクラスに所属するデータについて正しくそのクラスであると分類された割合を示す。Precision とは、あるクラスであると分類されたデータのうちどれだけが正しくそのクラスであるかという割合を示す。Recall と Precision は一般的には逆の相関関係にあり、識別器の性能が変わらない状態でひとつのクラスについての Recall を上げるということは、該当クラスであると分類するためのしきい値を下げることに相当し、誤判定が増えてしまうために Precision は下がってしまう。F 値とは Recall と Precision の調和平均のことであり、識別器の総合的な性能を示すために使われることが多い。研究では、新聞記事からプロフィール情報を抽出し、表として情報統合する実験を行っている。実験では2次の多項式カーネルを用い、多クラス問題の分類手法のひとつである pairwise 法を適用している。また、分割交差検定を行い、平均を最終結果としている。

山崎ら [84] は web 文書の読みやすさの研究に SVM を用いている。報告によると、文書の難しさや面白さなどの人の主観が強く働く評価に関する分類を行う研究は難しいために、分類手法についての報告が少ない。そこで、テキストの読みやすさをテキストの内容や語句の難易度で決めるのではなく、見た目や文の構造・論理構成で決まるものとし、テキスト構造の複雑さで読みやすさを評価する試みを行っている。研究では、修飾節と被修飾節の関係である係り受けが複雑であるほど日本語のテキストは読みづらくなるとして、係り受けの回数や節間の距離を特徴量として用いている。実験では、これらの特徴量をナイーブベイズや線形 SVM, 多項式カーネル SVM にて学習・識別させ、結果を比較している。実験の結果ではナイーブベイズに比べてどの SVM も識別性能が良かったとしているが、カーネルやパラメータのチューニングに関しては深く取り上げておらず、チューニングにより分類精度の違いを検討する必要があると述べている。

村上ら [85] は人物写真の顔画像における白飛びの検出に SVM を適用している。この報告によると、

画像検索において人物写真を証明写真などに利用する場合、顔写真を画像データベースから選定する作業は発生するが、候補画像が増えるほど選定作業量の増加が予想されるとしている。このとき、品質の良い顔写真を抽出することができれば、選定に関する作業コストを下げるができる。報告では顔写真の品質を決める要素として白飛びの定量的な評価を基準として選ぶことを提案しており、白飛びの検出に SVM を用いている。SVM に入力する特徴量としては、画像中の目や口唇などパーツ周りの白飛び率とし、顔写真について人間が白飛びの強弱を判断したものを正解ラベルとしている。SVM のカーネルには RBF カーネルを使用し、コストパラメータとカーネルパラメータはグリッドサーチを行って最適な値を求めている。ここで、グリッドサーチとは、パラメータの探索したい範囲と刻み幅を決め、それぞれのグリッドにおけるパラメータの組み合わせについて全数検査を行う手法である。グリッドサーチの特徴としては、勾配法などの探索法に比べて計算コストがかかるかわりに、局所解に陥りづらく、適切にグリッドが設定されていればパラメータの最適解が求まることである。実験結果によると、F 値で 90 % 以上の分類性能が得られたとしているが、肌の色が異なる場合は白飛びの発生量が変わる可能性があり、複数の肌色を基準とした SVM の学習を行う必要があるとしている。以上のように、SVM においては分類したいデータセットに合わせて適切に教師データの特徴量を選ぶ必要がある。

梶原ら [86] は疾病予防を目的として、生体情報と気象情報を用いた気分予測システムの開発を行っている。厚生労働省は近年の医療費の増加をうけ、生活習慣病や慢性疾患の予防を行うセルフメディケーションを推奨している。報告によると、ICT とセンサを用いてセルフメディケーションを支援する研究は盛んに行われており、疾病の早期発見、生活習慣等の改善による健康増進、疾病予測による予防の 3 つに分けられるとしている。報告では特に疾病の予測に関する研究を行っている。疾病は明確な症状が発生せずにユーザが「気分が悪い」など漠然とした表現しかできない場合もあるため、ユーザの健康状態を予測する手段として気分に着目したとしている。報告では、気分はそれを喚起した対象が不明確であることが多いが、「生理的な面に起因する体調の変化や外的要因に依存して緩やかに変化する持続性の高い感情」と定義している。そこで、気分の推定にユーザの生体情報、気象の環境情報、過去の気分を用いることで未来の気分を予測しようとしている。実験では、ユーザを 1 年間に渡って観測している。取得データは脈拍や基礎代謝量などの生体情報と、気温や降水量などの環境情報、その日における気分の尺度としている。実験では、機械学習の識別器としてナイーブベイズ、ランダムフォレスト、線形 SVM、RBF カーネル SVM を用いている。実験では未来の気分が Good か Bad かの 2 クラス問題を予測しており、特徴量として生体情報のみ、気象情報のみ、生体情報と気象

情報の3パターンを入力している。実験の結果では、気象情報のみを用いた識別結果は2クラスをランダムに選択した場合よりも悪かったとしており、気象情報のみから未来の気分を予測することは難しいと述べている。しかし、生体情報のみを用いた結果よりも、生体情報と気象情報を用いた結果の方が良くなっていることから、気象情報は何らかの要因を持ち、識別に有利に働いていると考えられる。

吉川ら [87] は SVM の利用法として、手の動作識別を筋電位から行っている。報告によると、義手を動作させる試みが行われている。パターン認識手法は筋繊維に発生する筋電位を利用して手の動作を推定するために用いられてきた。パターン認識が用いられる理由として以下の理由を挙げている。

1. 測定に用いる電極数よりも多くの動作数を識別できる。
2. 筋電位の個人差を吸収することで訓練の負担を減少させられる。
3. 個人内でも動作のたびに変動する筋電位から安定して動作識別できる。

そこで研究では SVM を手法として取り入れているが、SVM の特徴を次のように挙げている。

1. 大局的最適解が保証される。
2. 未観測のパターンに対しても識別能力が高い。
3. パラメータの探索が容易である。
4. 識別時の計算コストが少ない。
5. 学習を効率的に行うためのアルゴリズムが存在する。

報告では SVM の特徴を活かし、複数名の被験者による動作識別実験を行い、さらにリアルタイムに筋電位から動作識別をする実験を行っている。実験では SVM の比較手法として線形判別分析、k-最近傍法、バックプロパゲーションニューラルネットワークを取り入れている。SVM のカーネルには RBF カーネルを用いており、シグモイドカーネル、多項式カーネルと比較する予備実験を行った結果、RBF カーネルがもっとも識別性能が良好だったとしている。SVM のコストパラメータとカーネルパラメータはグリッドサーチを行っているが、予備実験を行ったあとのグリッドサーチでは、パラメータの最適な組み合わせと最も低い識別率となる組み合わせの差は6%であったが、最適な組み合わせから極端に離れたパラメータでない限りは、識別率に大きな低下はなかったと報告している。実験は多クラス問題であるため多クラスの識別手法として one-against-one 方式を用いている。Hsu ら [88] に

よる実験にて one-against-one 方式と one-against-all 方式を比べた結果、学習時間と識別精度に関して one-against-one 方式が優れているとの報告があるためとしている。吉川らの実験によると、SVM による動作識別結果は全被験者で 91 % 以上であり、被験者平均は 94.6 % であった。他の比較手法と比べても、SVM の結果が最も良好であった。また、筋電パターンの経時変化に対する実験においては、時間が進むにつれて識別精度が下がっていく傾向は見られるものの、90 分間に渡って 92 % 以上の識別率を維持している。吉川らによると、疲労や電極ずれなどにより筋電位の性質が変化する、つまり特徴量の分布が変化する状態においても、SVM は最初の学習によって経時変化をうまく識別できるモデルを獲得できているとしている。

SVM のカーネル開発は現在も続いており、新美ら [89] は時系列のデータストリームを入力とすることができるカーネルの研究を行っている。報告によると、ストリームカーネルは時間的な変化を特徴量にすることができるとしている。ストリームカーネルを用いた非線形 SVM は従来の非線形 SVM を上回る識別性能が得られたとしている。さらに、報告ではストリームカーネル SVM をパラレルブースティングとしてつなげることで、さらに性能を向上できるとしている。パラレルブースティングの特徴として、順応性の高さを挙げている。データストリームは学習モデルを常に更新しつづける必要がある。なぜならば、データストリームは時間的な変化を特徴量にしているためである。しかし、パラレルブースティングでは新たな教師データから構築した性能が低い識別器をアンサンブル学習モデルへ組み込むだけで更新できるとしている。アンサンブル学習とは、性能が良い識別器を構成するために、性能が低い識別器を組み合わせる手法である。また、通常のブースティングは性能が低い識別器をシリーズにつなげることで少しずつ重み付けをしていくが、パラレルブースティングは性能が低い識別器をパラレルにつなげることで並列計算ができ、計算時間を短縮することができるとしている。

三浦ら [90] は屋外移動ロボットの位置推定に関する研究に SVM を適用している。屋外での位置を取得する方法としては GPS があるが、ビルの密集地などのように高い建物の近くでは衛星が隠蔽されてしまい、位置情報の精度が下がってしまうという問題がある。また、測定した位置とその周りの環境との対応を示す地図を与える必要がある。三浦らの従来手法 [91] では画像特徴を手動で決定していたことを欠点として、提案手法では、ユーザがロボットを誘導して経路についての学習を行ったあと、ロボットが学習結果を用いてその経路上のどこにいるかを推定するという方法をとっている。位置推定の手法として、まず経路上で得られた画像の物体識別を行い、その識別結果を連結させて新たな特徴量を生成し、その特徴量を使って位置推定を行うことで、手動でのパラメータ調整を排除することができ、屋外での位置推定のロバスト性を向上させている。

千田ら [92] は RFID タグを用いた自己位置推定に SVM を適用している。報告によると、人の生活環境で活躍することが望まれるロボットには、移動の機能と自己位置推定の機能が必須である。さらに、自己位置推定に必要とされる精度は、ロボットの大きさや質量などに依存するが、仮に人間の半分から 2 倍程度の大きさだとすると、数 cm 程度の精度が必要であると述べている。既存の自己位置推定の方法は大きく分けて 2 通りあり、ひとつはタイヤの回転量計などを使って初期位置からの移動量を決める方法であるが誤差が累積するという問題がある。もうひとつは環境固定の目印をカメラ等で観測することで自己位置を推定する方法である。こちらの方法は累積誤差は発生しないが、目印が必ずしも観測できるわけではないとしている。そこで報告では環境固定の目印を使う方法として、RFID のタグに注目している。報告によると、RFID は安価でありすでに幅広い分野で活用されているシステムであるが、将来的には物流管理や道案内など様々な環境に設置されていくと予想しており、これらのロボット管理者以外が設置した RFID タグをロボットの自己位置推定に利用することで、ロボットの自己位置推定のためのインフラ整備にかかるコストを下げられるとしている。高精度の自己位置推定をするためには測定平面の大きさと分割数に応じた SVM の識別器が必要であるとしている。SVM はもともと 2 クラス問題を判別する識別器であるため、ロボットの位置を求めるためには測定平面のグリッド位置に対応する SVM の識別器を用意し、どの識別器が正の反応をしたかで位置を決める必要があるためである。また、グリッド状に配置する以外の方法として、SVM の識別器を木構造で並べることで二分探索で位置を決める方法も提案している。グリッド状に配置する方法に比べ、木構造に配置する方法では識別器の数が 2 倍必要であるが、計算コストは低い。

小中 [93] は離散値入力制御系に対して SVM を適用し、制御器を設計する手法について研究を行っている。電力変換や化学プラントなど様々な分野でスイッチング制御が使われており、これは入力をスイッチという離散値に制限した制御系としてモデル化される。小中によると、このような制御系を設計する問題は、観測出力と離散値入力において離散値入力を切り替える条件問題だとし、問題に機械学習を取り入れることで期待される効果としては以下の 3 点を挙げている。

1. 制御対象の内部モデルを構築する必要がなく入出力の時系列変化のみ取得できればよい。
2. 制御対象が線形であっても非線形であっても適用できる。
3. 非線形な制御器を自動的に構築できる。

この報告によると、制御器の設計に機械学習を取り入れる先行研究としてニューラルネットワークを用いたものがあるが、ニューラルネットワークの学習効果は初期の重み付けや教師データの選択順序

によって変化してしまうため、制御対象を実際に動かせる機会が限られる場合には適切でないとしている。そこで SVM を取り入れるメリットとしては以下の理由を挙げている。

1. 教師データが同じならば識別器が確定的である。
2. パラメータが少数である。
3. 追従性能が高い。

SVM に入力する特徴量として、制御器の入力（スイッチの状態）・出力・目標値との偏差を時系列データとして与え、目標値に近づけるためのスイッチの状態を SVM からの出力としている。実験では、変動量の大きさによって 5 状態に分けられる入力をもつサーボモータの出力角度を SVM と比較手法である近似最近傍法の 2 つで制御する試みを行っている。実験の結果では、SVM は近似最近傍法よりも追従性がよく、PWM 信号のような制御入力選ばれている。実験における SVM の適用において、モータ内部のパラメータは一切使用していないことを小中は強調しており、SVM を適用するにあたり内部モデルのパラメータを考慮する必要がないことの例となる。また、教師データの数について、SVM は近似最近傍法の半数しか与えられていない。理由について報告の中では特に言及していないが、少ない教師データでも十分な識別器が作成できることを示唆している。

Munder ら [94] は画像中に歩行者が存在するかどうかの研究に SVM を用いている。この報告によれば、画像中の歩行者認識は調査研究やロボティクス分野など幅広い分野で重要な技術であり、人のポーズや服装などが多岐にわたるため、コンピュータビジョンの分野では非常に難しい分野になっている。報告の中で、SVM 以外に k-最近傍法やニューラルネットワークを比較対象として用いている。実験結果によると比較手法よりも SVM の識別精度が良かったとしており、さらに、教師データ数とテストデータ数の比を 1:3 から 2:3, 4:3 と教師データを増やすように変えていったところ、識別精度は向上していき、4:3 の比においても精度向上が飽和していないことを述べている。

以上のように、SVM はパターン認識が適用できる分野において広く使われている。

Cortes ら [69] によって考案された SVM は、2 クラス問題のための機械学習器として開発されているが、Chung ら [82] による SVM ライブラリの開発や、山崎ら [71] や千田ら [92] によるマルチクラス問題への適用や、新美ら [89] によるパラレルブースティングの適用など、現在も開発が進んでいる。SVM は山崎ら [84] や梶原ら [86] が報告しているように、ベイズ推定よりも識別性能が高い。

伊波ら [70] がシステムコールの異常検知に SVM を適用している例で述べているように、False Positive の減少に有効である。猿の出現予測でいう False Positive とは、実際には猿の群れがない状

態にも関わらず、猿が出現すると予測してしまうことに相当し、出現であると判別してしまうと住民は猿の襲来に備えて追い払いの準備をする必要がある。このような誤判別が続くと、次第に住民は予測システムを信用しなくなっていくと考えられるため、False Positiveの減少は猿の出現予測にとって非常に重要な意味を持つ。

小林ら [73] が述べているように、SVM に使用する RBF カーネルでは、コストパラメータとカーネルパラメータの調整だけで識別の汎化能力を最大化することができる。小林らはパラメータの決定にあたり勾配法では局所解に陥る可能性があるとして強化学習を取り入れ、パラメータ決定までの計算時間を削減する試みを行っているが、計算時間が問題になるのは学習と識別をオンタイムで行い続ける場合のみである。通常は、学習のために事前に十分な計算があるため、村上ら [85] が述べているようにパラメータのグリッドサーチをすればよく、グリッドサーチの場合は勾配法のように局所解に陥らずに最適解を得ることができる。本論でも SVM のパラメータ決定にはグリッドサーチを用いる。

ベイズ推定では特徴量の選択が識別性能の向上には欠かせなかった。SVM についても青木ら [74] は特徴量の次元数を削減する試みを行っているが、特徴次元数が数十個から成るデータを対象としており、また、特徴次元数が数十次元から十数次元の間ではほとんど識別性能は変化せず、5次元よりも少なくすると、今度は識別性能が劇的に劣化することも示している。特徴次元数による SVM の識別性能の変化がほとんどないことは Melgani [79] からも報告しており、Melgani の実験結果では特徴次元数が 20 から 200 まで変化させたとしても識別性能の分散が 0.3 であったとしており、次元数が増えたとしても SVM は識別平面の決定時に意味のある特徴量を自動で用いるので、精度の変化はほとんどない。本論では猿の出現予測を行うにあたって環境要因として 8次元を選択しており、8次元よりも次元数を削減する必要はないと考えられる。また、環境要因として次元数を増やすとしても、特別な観測をせずに得られる環境要因の数は限られており、本論が最終目標としているのは、猿専用特別な観測体制を敷かなくても猿の出現予測を行えることであるため、本論の特徴量次元数として環境要因の 8次元は適切なサイズであると考えられる。

吉川ら [87] は SVM で筋電位から手の動作識別を行う研究を行っているが、SVM を取り入れた理由として、上述の最適解の保証やパラメータの探索の容易さに加え、未観測のパターンに対しても識別能力が高いことを理由として挙げている。ベイズ推定は確率モデルに基づく予測手法であることは 2.3 節ですでに述べたが、猿の群れの長期的な観測は難しいため、観測済みのパターンによる確率モデルが必要なベイズ推定は高い識別性能が得られない。そこで、未観測のパターンに対しても識別能力が高い SVM を手法として取り入れることは本論にとって非常に重要である。また、Munder ら [94]

が教師データ数を増やす実験を行った結果では、教師データが増えていっても識別精度の向上は飽和しないという点も本論にとっては重要であり、未観測のパターンが多い状態から、観測済みのパターンが増えていく状態まで対応可能な SVM は有効な手法である。

また、2.2 節でも述べたように、猿の群れの行動に対して内部状態の推定を行うことは困難である。猿の群れの内部状態とは、なぜ移動したのか・群れの意思決定はどのように行われているのかなどの、表層的にあらわれる状態を決定するための要因であり、人間が猿の群れの内部状態を推定するためには長期に渡って群れの社会学的な調査も含めて行う必要があり、現実的には難しい。そこで、小中 [93] が述べるように、内部状態を推定するのではなく、入出力を与えるだけで対象のモデル化ができるという特徴が、猿の群れの行動予測には必要である。また、教師データを与える順番などによって学習効果が変わってしまうニューラルネットワークのような識別器よりも、SVM のように同じ教師データであれば同じ学習効果が得られる識別器を用いる必要がある。猿の群れの行動予測を行うにあたって、学習の順番によって識別結果が変わってしまうということは住民からの信頼性の低下を招く恐れがあるため、学習効果が同一である保証が必要となる。

梶原ら [86] は人間の気分を予測する研究に SVM を適用している。気温や降水量のような環境要因を特徴量に取り入れることで人間の気分をモデル化することができるというのは本論にとっても非常に有効な知見である。本論は猿の出現予測を行うにあたって、同様に環境要因を取り入れているが、猿も生物である以上、環境要因に行動を依存すると考えられる。例えば、非常に寒い時期や降雨の激しい時は体温低下によるカロリー消費を防ぐために行動を控えたり、また、逆にカロリー取得を優先して摂食行動が激しくなる可能性が考えられる。直接的に環境要因が猿に影響を与えることも考えられれば、2.1 節で述べたように、人間の活動を避けるという意味で環境要因の従属変数になっている可能性もある。つまり、降雨の激しいときは人間が野外での活動を控えると考えられるので、人間を避けるような行動をとる猿にとっては、摂食しやすい環境になっているともいえる。

以上のように、猿の行動推定に適していることから、本論では猿の出現予測に SVM を採用する。

2.5 本論の関連研究

これまでに猿を始めとした獣害対策として、様々な研究を行っている。本節では、それらの代表的な物の説明を行う。

2.5.1 項では、RSSI を利用した進行方向の推定に関する研究 [95] の説明を行う。

2.5.2 項では、RSSI を利用した位置推定に関する研究 [96] の説明を行う。

2.5.3 項では、赤外線情報を利用した害獣の計測に関する研究 [97] の説明を行う。

2.5.4 項では、猿の襲来に関する分析研究 [98] の説明を行う。

2.5.1 RSSI を利用した進行方向の推定

RSSI の閾値を決めるにあたり、電波の減衰特性から距離を求める手法は存在する。Whitehouse ら [99] は屋内や屋外において RSSI から位置を推定する試みを行っているが、どちらの環境においても開けた平地となっている。趙ら [100] は小型飛行ロボットを使って屋内を走査して RSSI を測り、電波地図を作る研究を行っているが、本論では屋外の広域な地域を対象としているので、全ての地域で RSSI を測るというのは現実的ではない。日坂ら [101] は市街地において RSSI を用いて交差点での歩行者検出をしているが、市街地であることに加えて複数のアンテナを近接させて用いているため、山間地向きではない。立石ら [102] はシミュレーションによって RSSI の減衰係数を推定しているが、屋内を模擬しているため、屋外においてどのような結果になるか不明である。

以上の研究は全て 2.4GHz 帯の無線機を想定しており、本論のように 150MHz 帯で山間部を対象とした研究は少ない。特に本論のような実験地は地形の変化が激しく、理論的な距離への変換式を成り立たせるのは難しい。そこで中山ら [95] は簡易にキャリブレーションを行う手法を開発した。

提案するキャリブレーション手法では、受信機で受信した RSSI の最高値と平均値を使って距離換算用の補正係数を求めることにより、アンテナやケーブルでの損失や、受信機の感度をひとつひとつ較正する必要がなくなる。また、受信機の設置環境に応じた距離換算用の補正係数が得られるため、山間部における電波伝搬特性の違いを吸収することができる。

キャリブレーション手法による距離補正の精度向上結果を示し、さらに猿の進行ルートを推定する実験も行っている。

2.5.2 RSSI を利用した位置推定

西岡ら [96] は、山間地と隣接する小規模な市街地にて発信機の RSSI を利用して猿の位置推定をする研究を行っている。研究では、受信可能範囲が重なるように受信機を配置し、事前準備として発信機と GPS を持ち歩いて RSSI と位置情報をセットで記録していき、RSSI の強度地図を作成する。RSSI から位置を推定するとき、それぞれの受信機の強度地図から RSSI 生成確率を計算し、受信機毎の生成確率地図をかけ合わせることで位置を推定している。さらに、移動可能な範囲を考慮することで、移動経路を絞り込んでいく処理を含めている。

以上のような FootPrint を利用する手法は、マルチパスなどによる RSSI の変動も全て記録しておくことで位置推定の精度を上げられるが、受信機をオーバーラップするように配置するためコストが高いことや、事前に発信機と GPS を持ち歩いて推定をかけたいエリアの RSSI と位置情報を全て記録しておく必要があるなど山間地には適用しづらい手法となっている。

2.5.3 赤外線情報を利用した害獣の計測

北原ら [97] は、害獣対策の一環として、害獣をカウントする研究を行っている。害獣を罠で効率良く捕獲するためには、罠に害獣が可能な限り多く入ったタイミングで罠を起動する必要があるが、罠で捕獲可能なシカやイノシシなどの害獣は夜間によく活動する。そのため、夜間に罠を監視し続けるのは人間にとって負担が大きく、さらに可視光の投光器を使ってしまうと害獣が寄り付かなくなるという問題がある。そこで、赤外線により深度情報を画像として得られるカメラを使い、害獣の写った深度画像に背景差分、二値化、分散値による切り分けなどを行うことで害獣を 1 頭ずつ切り分け、カウントを行っている。

2.5.4 猿の襲来に関する分析

伊藤ら [98] は、猿の襲来規模に関する研究を行っている。猿の出現を長期的に観測した結果から、猿の群れの中にも複数のグループがあることを示し、複数のグループで同時に出現するときにキーとなるグループ、つまり他のグループとよく同時に出現するグループがあることを発見した。このキーとなるグループの動きを捉えることができれば、猿の出現を精度良く予測できる可能性を示している。

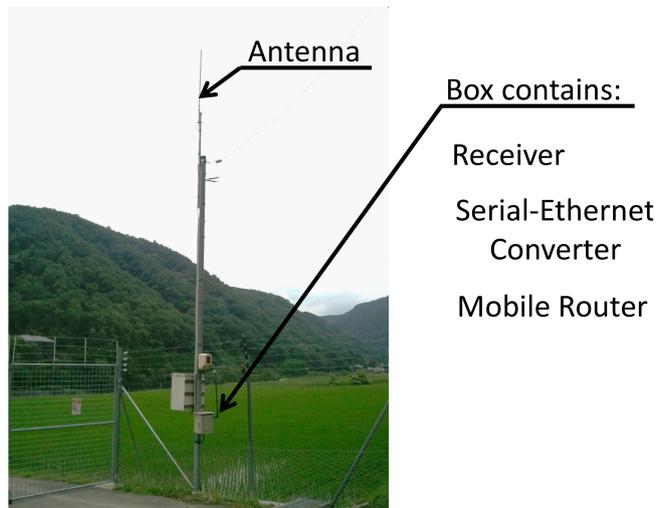


図 2.1: アクセスポイントの一例
Fig. 2.1 An example of access points.

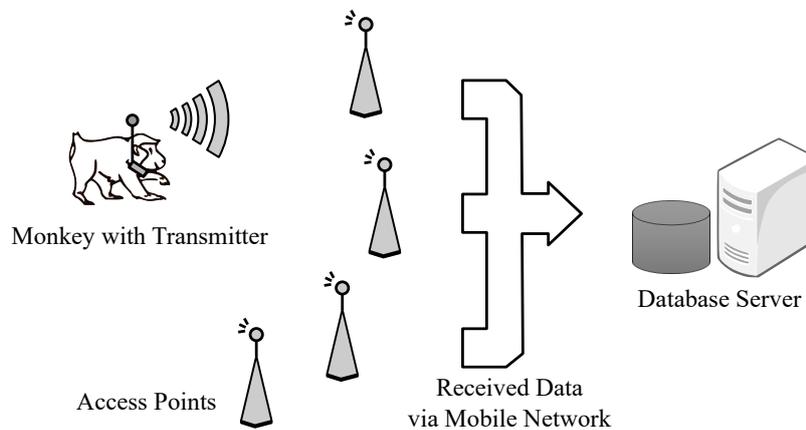


図 2.2: システムモデル
Fig. 2.2 System model

2.6 実験システムの構成

2.6.1 システム設計

猿の出現予測を行うにあたり、こういった条件のときに猿が農作地に接近するのか調べる必要がある。そこで、猿に発信機付きの首輪を取り付け、集落に図 2.1 のような受信機（アクセスポイント: AP）を設置し、猿の接近を検知するシステムを構築した。AP は概ね集落の中心に設置した。AP は発信機の電波を受信した際に、日時、発信機番号、受信機番号、受信信号強度（RSSI）をサーバに記録する。システムモデルを図 2.2 に示す。電波は RSSI の強さに関わらず、受信したものは全て記録する。

AP を三重県伊賀市の猿が出没する山を囲むように、集落沿いに 21ヶ所設置した。AP の設置場所

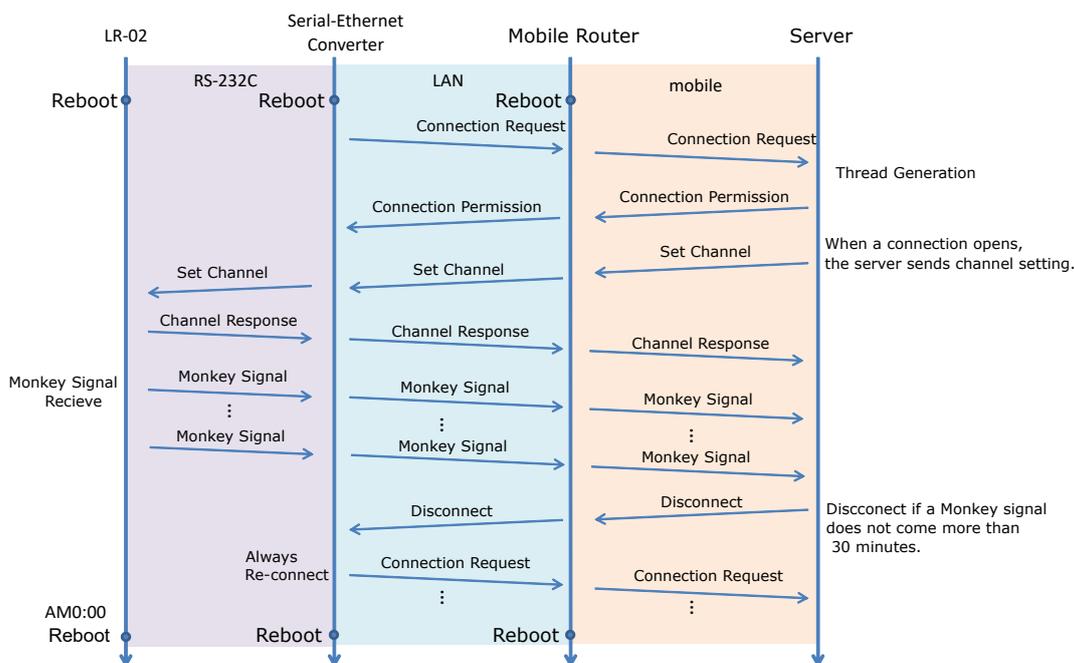


図 2.3: LR-02 からサーバまでのデータフロー
Fig. 2.3 Data flow through LR-02 to the server.

は図 2.4 に二重丸で示す場所である。AP 間の距離は 1~2km に設定しており、猿がどの経路で山から農作地へ近づいても捕捉できるように意図している。つまり、農作地（猿にとっての良質な餌場）に近づくほど RSSI が強く記録される。

2.6.2 猿の行動収集

発信機は、ある群れの雌猿 7 匹に取り付けた。群れは 100 匹程度の群れであると推定されている。猿は母系の小集団で行動することが報告されており [103]，群れ全体の 10 % 程度の頭数に発信機をとりつけば、群れ全体の動きを捉えることができると考えられる。

猿に取り付けた発信機は動物検知用の送信機の LT-01 である。LT-01 は 150MHz 帯の特定小電力無線機器であり、標準規格 (ARIB STD-T99) の技術基準について、適合証明を取得している。発信機の諸元は付録の表 A.1 に示す。

AP に使用した受信機は、LT-01 と組み合わせて使用できる動物検知用の受信機の LR-02 である。LT-01 と LR-02 の組み合わせでは 1km 程度の距離での通信が可能である。受信機の諸元は付録の表 A.2 に示す。

送信機と受信機の特性を鑑みた結果として、AP 間の距離は 1~2km に設定している。LR-02 にはシリアルイーサネットコンバータと携帯通信回線をもつモバイルルータを接続し、サーバにデータ

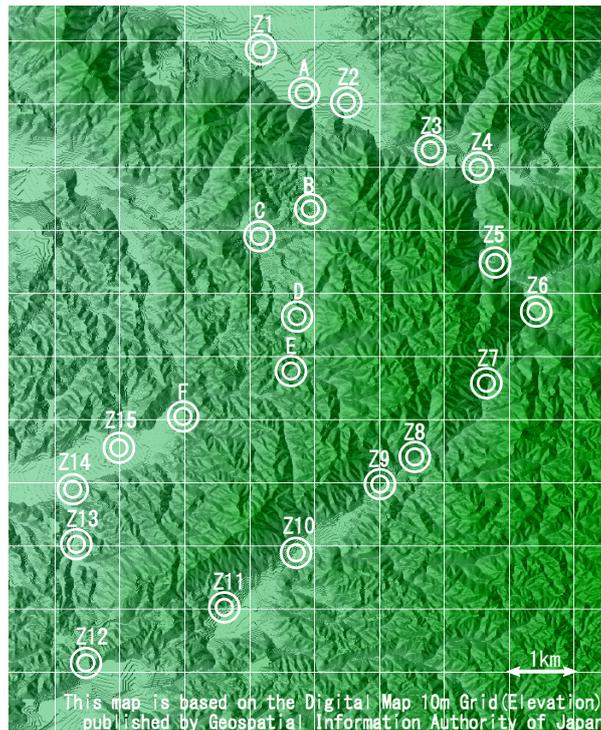


図 2.4: アクセスポイントの配置 (二重丸がアクセスポイントの位置を示す)
 Fig. 2.4 Access point locations. The double-circles indicate the access points.

を送信できるようにした。図 2.3 に、LR-02 からサーバにデータを送るまでにどのようなプロトコルになっているかを示す。AP は 2012 年 6 月に設置を完了した。同年 6 月中は安定稼働させるための準備期間とし、正式な記録は 2012 年 7 月 1 日から 2014 年 5 月 15 日までの記録とした。AP の設置前に猿への発信機の取り付けは行われており、記録を 2014 年 5 月 15 日で終了した理由は、発信機の電池が十分に持つ期間内を正式な記録とするためである。終了日以降も発信機の電波は受信していたが、徐々に受信回数が減っていき、最終的には受信が途絶えた。

2.6.3 AP の稼働率

猿の信号を受信していない時、AP は 30 分毎に通信確認をサーバに行い、AP が生存していることを記録した。ここでいう稼働率とは、対象期間において LR-02 からサーバまで正常にデータを送信できる状態であった時間の割合のことである。サーバまで正常にデータを送信できる状態の意味は、猿の信号を受信している時間と、通信確認による生存している時間をさす。2012 年 7 月 1 日から 2014 年 5 月 15 日まで全 AP での稼働率は 97.22 % となった。システムの要素が直列に接続され、どこかの要素が止まるとシステム全体が止まってしまうような直列システムの稼働率は、それぞれの要素の稼働率（信頼性）の積をとることで表現できる。本システムの要素としては以下の 5 つがある。

1. 電源供給
2. 受信機
3. シリアル-イーサネットコンバータ
4. モバイルルータ
5. サーバ

仮にそれぞれの稼働率が99.5%だとすると、 $0.995^5 = 0.975$ となり、本システムの稼働率である97.22%と近い値になる。実際には耐候性の観点からそれぞれの機器の稼働率は下がると考えられるため、本システムでは良好な稼働率が得られたと考えられる。稼働率が下がった主な要因として、電源が抜かれた等の人的要因、3G回線の接続不良が挙げられる。特に稼働率が低いAP番号Z1とAP番号Dは理由のほとんどが電源の切断であり、再度電源を入れるだけで復旧できた。

表 2.1: アクセスポイントの稼働率

Table 2.1 Availability of Access Points

AP	Availability (%)	AP	Availability (%)
A	97.15	Z6	98.89
B	97.02	Z7	98.54
C	94.54	Z8	96.23
D	93.94	Z9	99.50
E	94.07	Z10	97.29
F	98.44	Z11	97.78
Z1	91.18	Z12	98.29
Z2	99.36	Z13	98.84
Z3	97.34	Z14	98.55
Z4	98.63	Z15	98.18
Z5	97.89		
		Average	97.22

2.6.4 受信信号強度による接近検知

2012年7月1日から2014年5月15日までの猿の発信機からの受信回数は約2500万件であった。受信結果の中には、猿が集落から遠い場所にいる（つまりRSSIが低い）状態を含む。猿は捕食活動が終わった後に山に帰らず、農場から少し離れた茂みに隠れることが目撃されており、今回の実験で得られた結果であるRSSIが低い状態が多いことがそれを裏付けている。そこで、猿が充分接近した

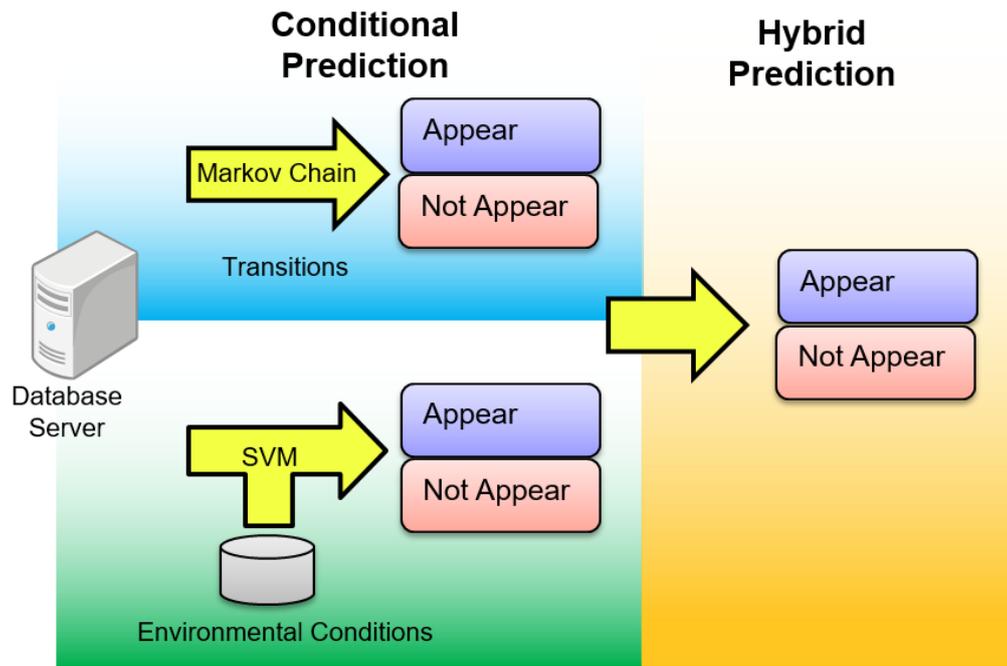


図 2.5: 出現の予測方法
 Fig. 2.5 Methods for Appearance Prediction.

と判断するために RSSI の閾値を決める必要があるが、猿の出現の予測手法にも関連するため 3.2.2 項で説明する。

2.7 予測システムの構成

2.6 節で説明してきたように、本論では猿の行動を収集する実験システムを構築した。実験システムによって得られた2年間の猿の行動データを用い、3章、4章で猿の出現予測を行っていく。図 2.5 に示すように、猿の実験データをもとに予測を行っていく。

2.1 節にまとめたように、猿は行動圏を持ち、農作地や山を餌場としながら移動している。猿は生物の中でも比較的人間に近い知能を持っていると考えられるため、人間の行動モデルと同様の手法が適用できる可能性がある。そこで2.2 節で述べた、人間の行動モデルの推定によく適用されるマルコフチェーンモデルを用いて猿の移動をモデル化していく実験を3章にまとめる。これにより、出現のパターンから1日単位の出現を予測する。設置したAP間の距離に対して、猿の移動速度には上限があるため、多数のAPに同じ日に出現するという事は少ない。そこで、1日を単位として猿がどこのAPに出現したかのデータをもとにマルコフチェーンで猿の移動をモデル化し、前日までにどこのAPに猿が出現していたかという情報が予測の条件として与えられるとき、猿が次に出現するAPの予測を行う。

猿の群れの行動は非常に非線形な挙動をすると考えられるが、2.1 節に示したように、環境の状態によって行動が変化する可能性がある。そこで、2.4 節でまとめたSVMを手法として用い、環境要因によって猿の行動をモデル化する試みを4章にまとめる。2.3 節のベイズ推定は確率的なモデルであり、猿の行動の推定には向かないと予想されるため、SVMの比較手法として使用する。4章では出現予測の単位を1時間のスロットとして行う。猿の行動は環境の状態によって変化する可能性があるが、環境要因自体は猿がどこのAPに出現するかを説明するものではない。4章での出現予測は、ある時間に猿がどこかの集落に出現するかしないかの2クラス問題を取り扱う。

4章の最後では、移動要因と環境要因を組み合わせた出現予測を本論の提案手法としてまとめる。移動要因は1日単位の予測、環境要因は1時間単位の予測であり、それぞれ独立して精度を検討できるが、2つの手法を組み合わせることで1時間単位の出現予測の精度を向上させられる可能性がある。

第3章 移動要因を用いた猿の出現予測

3.1 緒言

本章では，発信機を最終的には使わないという条件を満たしつつ猿の出現パターンを直接的にモデル化するために，マルコフチェーンを手法として用い，出現予測を行うことを目的とする。

猿は餌を求めて農作地や山を移動していく。農作業者のように個別の農作地のみを観測しただけでは猿がどのようなパターンで移動していくかは判断しづらい。そこで，まずは猿がどのように移動しているかを移動地域全体に渡って把握する必要があるため，行動収集用の実験システムの設置を行った。

実験システムによって得られた移動データをもとに，猿の出現数が多い農作地を選択し，猿が1日毎にどのように農作地や山を移動しているかをマルコフチェーンによりモデル化する。マルコフの状態遷移確率を基に猿が出現するかどうかの予測を行う。

本研究の予測精度として猿が出現するかどうかの2クラス問題で80%以上，どの集落に出現するかマルチクラス問題で50%以上にすることを最終的な目標としている。目標値の参考として，人間の行動をマルチクラス問題で扱った研究 [25][31][32] で予測精度が50%前後であることから，本論のマルチクラス問題も50%以上とし，2クラス問題はマルチクラス問題よりも精度が上がると想定されるため，80%とした。

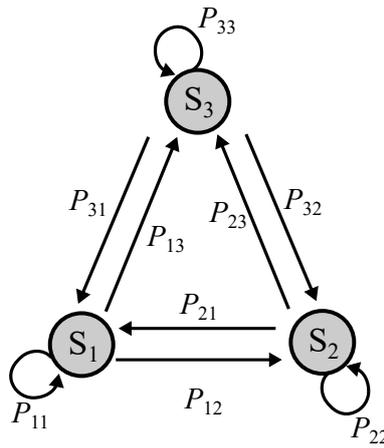


図 3.1: マルコフチェインの基本概念
 Fig. 3.1 Basic concept of Markov Chain.

3.2 原理

3.2.1 マルコフモデル

ある時間 t における状態 $S_t (t = 1, 2, \dots, n)$ を考えたとき, S_t は直前の状態 S_{t-1} によって決まるとする. このとき, S_{t-1} を i , S_t を j とすると, 状態遷移の確率は P_{ij} とあらわすことができ, P_{ij} は t には依存しない. このとき, S がとりうる状態の数を C とすると, 以下の式が成り立つ.

$$\sum_{j=1}^C P_{ij} = 1 \quad (i = 1, 2, \dots, C) \quad (3.1)$$

以上のように, 現在の状態が直前の状態によって決まる性質を持つ事象をマルコフ過程といい, とりうる状態が離散的なものを特にマルコフチェインという. さらに, 現在の状態が直前の M 時点前からの状態によって決まるものを M 重マルコフチェインという. 本稿では, $M = 1$ のものを単純マルコフチェインと呼ぶ. Fig. 3.1 は, 状態数 $C = 3$ としたときの単純マルコフチェインの状態遷移図である.

3.2.2 猿の出現の定義

関連研究で開発されたキャリブレーション手法により距離と RSSI の対応テーブルを作成した. キャリブレーションテーブルと RSSI を見比べ, 猿の接近を判断する. 集落の立地などを鑑み, 推定距離が 150m 以内を猿の接近とした. 1 日のうちに 1 回でも接近があればその AP に出現 (Appearance),

表 3.1: それぞれのアクセスポイントに猿が出現した日数
 Table 3.1 Number of days which monkeys appear at each access point.

AP No.	1st year	2nd year
N	76	81
A	72	22
C	42	51
D	42	45
F	30	49
B	22	28
E	5	18
Z	23	71
Total	312	365

1 回も接近がない日は非出現 (Non-appearance) と定義した。同日に複数の AP で接近があった場合は、受信回数が多い AP に出現とした。記録期間から停電などでシステムが停止した期間を除いた日数は 677 日であった。また、3.3 節のマルコフチェーンモデルの教師データ、予測対象データとするために、1 年目を 312 日、2 年目を 365 日として区切った。

実験期間中にどこの AP で何日分の出現があったかを Table 3.1 にまとめた。Table 3.1 は、AP 番号ごとに日数が上位のものから記載した (AP 番号 Z は降順ソートから除く)。

AP 番号 N は非出現をあらわす番号とした。AP 番号 Z は表 3.1 にない AP 番号をその他の AP としてまとめた番号である。図 3.2 は、猿がどのように AP を移動していったかをあらわしたものである。横軸が経過日数、縦軸がどの AP に出現したかを示す。図 3.2(a) は 1 年目、(b) は 2 年目の移動をまとめたものである。縦軸は AP 間の位置関係を考慮しており、AP 番号 N と Z 以外は図 2.4 に示した AP の位置関係の通りに上下に並べる。

図 3.2(a) の 120 日目付近までと図 3.2(b) の 270 日目付近までは春から秋に相当するが、活発に移動している様子が見て取れる。原因として、春から秋は山や農作地に食料が豊富なため、どちらの餌場も利用していることが考えられる。また、農作地に出現することによって人との遭遇リスクを減らしている可能性も存在する。図 3.2(a) の 120 日目付近から 210 日目付近にかけては AP 番号 A に集中して出現している。原因としては、山に食料が少ないために食料を得やすい集落に出現していることが考えられる。図 3.2(a) の 210 日目付近から 310 日目付近にかけては冬から春に相当する。この時期は猿の移動が増え始めるが、1ヶ所での出現日数は長めになっている。理由として、AP 番号 A の食料が尽きた可能性や、できるだけ 1ヶ所で長く採食することで、移動によるカロリー消費を抑えている可能性もある。図 3.2(a) と (b) を比較すると、同じ時期では同様の傾向がある。

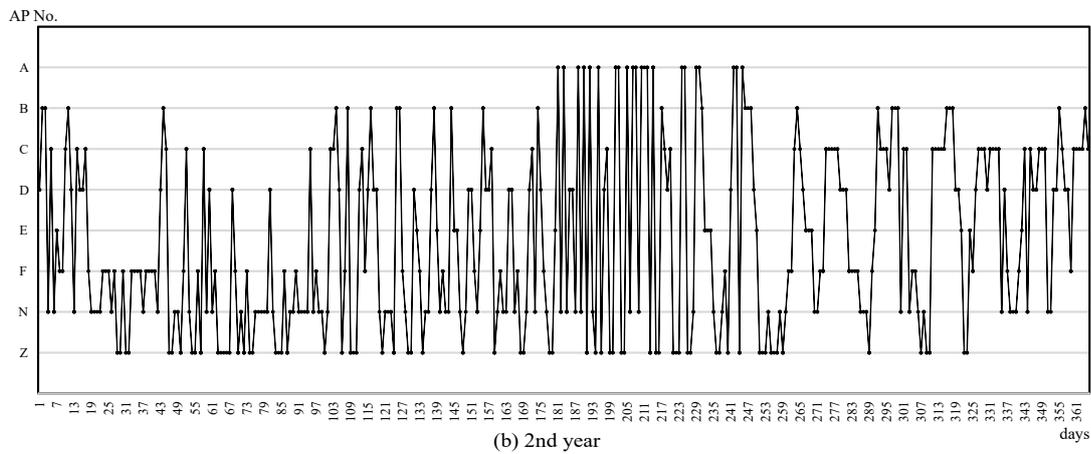
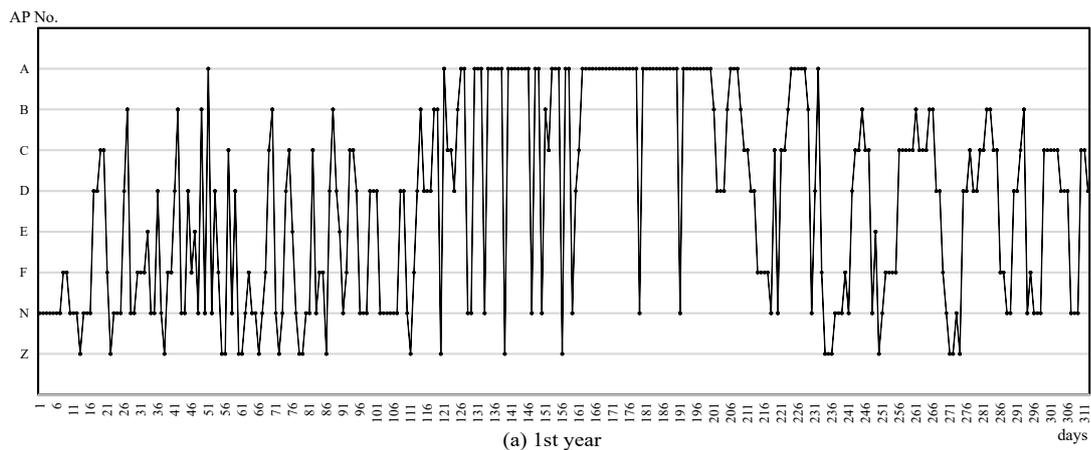


図 3.2: アクセスポイント間における猿の遷移活動
 Fig. 3.2 Monkeys' transition activity through each access point.

3.3 実験方法

3.2.2 項で、ある日に猿がどの AP に接近したかということを出現と定義した。マルコフチェーンで扱う状態は、この定義での AP 番号である。マルコフチェーンの状態遷移確率は、実験期間の 1 年目を使って計算した。1 年目の実験日数の合計は前述の通り 312 日であった。状態遷移確率は単純マルコフから 5 重マルコフまでを計算した。表 3.2 に単純マルコフチェーンの状態遷移確率を示す（2 重以上の結果は行数が長くなるため省略する）。Before はマルコフチェーンの遷移前の状態、After は遷移後の状態をあらわす。

図 3.3 は表 3.2 の単純マルコフチェーンの状態遷移の中から、それぞれの遷移確率の上位 2 つを抜き出し、AP の位置関係を保ったまま遷移図としてあらわしたものである。AP 番号 N と Z は位置関係を明確に決めることができないため、便宜的に図上に置いた。

1 年目の状態遷移確率を教師データとしてマルコフチェーンの分類器を作成し、2 年目に猿がどのように状態遷移していくかを予測する。1 つの遷移前の状態に対して、次にどの状態に遷移するかは、遷移確率が最大の状態に遷移することとした。遷移確率が同じ場合は、1 年目の出現日数が多い AP に遷移することとした。以上の条件から、遷移確率がゼロの場合は AP 番号 N に遷移すると予測することになる。また、分類器で予測対象とする日 (t_n) より以前の時刻 ($t_{n-m} \cdots t_{n-1}$) の状態は実験による正解データを与えた。遷移前の状態として正解データを与える理由は、2 年目の 365 日分の猿の予測を行うにあたって、予測結果のみを使って予測していくとすぐにマルコフ過程が吸収状態に陥ってしまうため、吸収状態を避ける意味もあり、また、実際のシステムの運用を行う場合でも前日までに猿がどの AP に出現したかという情報は共有しやすいためである。

予測結果は (3.2)-(3.5) 式に示す指標を計算して評価した。あるクラス c_i の予測を行ったときに、正しくそのクラスであると分類されたサンプルを True Positive (TP), c_i 以外のクラスであると分類されたサンプルを False Negative (FN) と呼ぶ。また、別のクラス c_n の予測を行ったときに c_i であると分類されたサンプルを c_i にとっての False Positive (FP), c_i 以外であると分類されたサンプルを True Negative (TN) と呼ぶ。(3.2)-(3.5) 式に示す指標はこれら TP, FP, FN, TN を用いて以下のように計算する。

$$Precision = \frac{TP}{TP + FP} \quad (3.2)$$

表 3.2: 単純マルコフチェーンの遷移確率

Table. 3.2 Transition probability of 1st order Markov Chain.

Before\After	N	A	B	C	D	E	F	Z
N	0.421	0.079	0.026	0.066	0.171	0.013	0.118	0.105
A	0.111	0.792	0.042	0.014	0.000	0.000	0.014	0.028
B	0.273	0.136	0.136	0.227	0.182	0.000	0.000	0.045
C	0.095	0.048	0.167	0.476	0.143	0.024	0.048	0.000
D	0.122	0.024	0.171	0.171	0.366	0.024	0.098	0.024
E	0.800	0.000	0.000	0.000	0.000	0.000	0.000	0.200
F	0.233	0.000	0.000	0.100	0.067	0.067	0.400	0.133
Z	0.391	0.130	0.000	0.043	0.087	0.000	0.087	0.261

$$Recall = \frac{TP}{TP + FN} \quad (3.3)$$

$$F - measure = \frac{2Precision \times Recall}{Precision + Recall} \quad (3.4)$$

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \quad (3.5)$$

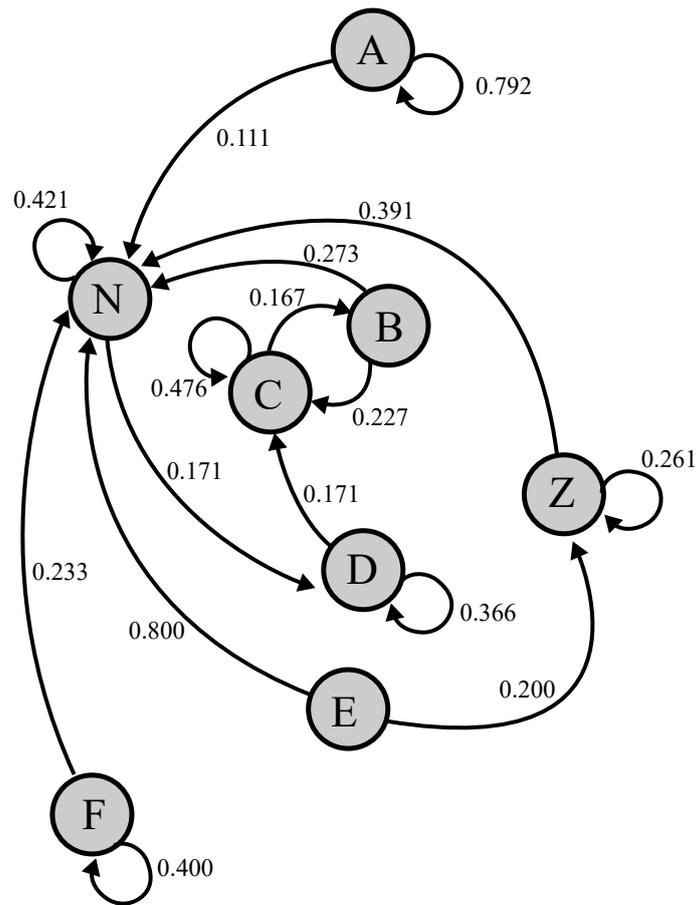


図 3.3: アクセスポイントの配置と単純マルコフチェーンの代表的な遷移確率
 Fig. 3.3 Access point locations and representative transition probabilities of 1st order Markov chain.

表 3.3: 2 クラス問題の結果

Table 3.3 Result of two-class problem

Evaluation	Class	1st	2nd	3rd	4th	5th
Precision	Non-appearance	0.232	0.280	0.220	0.213	0.218
	Appearance	0.790	0.827	0.775	0.743	0.735
Recall	Non-appearance	0.568	0.580	0.605	0.765	0.889
	Appearance	0.465	0.574	0.387	0.194	0.088
F-measure	Non-appearance	0.330	0.378	0.322	0.333	0.350
	Appearance	0.585	0.678	0.516	0.307	0.157
Accuracy	-	0.488	0.575	0.436	0.321	0.266

3.4 出現/非出現の2クラス問題の結果

表 3.3 は猿がどこかの AP に出現 (Appearance) するか, それとも非出現 (Non-appearance) になるかの 2 クラス問題として予測を行った結果である. つまり, 正解データと予測結果が共に AP 番号 N であった場合は非出現として正解, 正解データと予測結果が共に AP 番号 N 以外であった場合は出現として正解, それら以外の場合は不正解とした. 5 次までの M 重マルコフチェーンを予測に使用した結果, 2 重マルコフチェーンの結果が最も良く Accuracy で 57.5 % となり, 2 重より大きい次数のマルコフチェーンでは Accuracy が下がっていく結果となった. 猿の追い払いをしなければ作物が食べられてしまうため, 猿が出現するかどうかは農家にとって非常に切実な問題である. 57.5 % では少し心もとないが, 半分以上の割合で予測が当たるのであれば農家にとっても十分に有益な結果といえる. 表 3.3 の結果は, ほとんど全ての項目において 2 重マルコフチェーンが良くなっているが, 非出現の Recall だけは 5 重マルコフチェーンが最も高くなっている. マルコフチェーンの次数 M が増えるに従って, AP のとりうる状態の数が S_n^M で増えていくが, 今回の教師データの数が高次のマルコフチェーンの状態の数に対して少ない. そのため, 教師データに占める個数が多い AP 番号 N (非出現) への遷移が多くなり, AP 番号 N と判定される誤検出が増えたことにより非出現の Recall のみが高くなっていく結果となっただけである. 本研究がターゲットとする猿の獣害対策においては, 非出現の Recall だけが高い, つまり本来は猿が出現するにも関わらず非出現であると農家に伝えてしまうと猿の被害を見過ごすことにつながってしまうため, 5 重マルコフチェーンの結果は採用しない.

3.5 出現 AP のマルチクラス問題の結果

表 3.4 は猿がどこの AP に出現するかのマルチクラス問題として予測を行った結果である。つまり、正解データと予測結果が同じであった場合に正解、それ以外の場合を不正解とした。AP 番号 N（非出現）の結果は表 3.3 と同じである。2 クラス問題と同じように 5 次までの M 重マルコフチェーンを予測に使用した結果、2 重マルコフチェーンの結果が最も良く Accuracy で 31.5 % となり、3 次から 5 次までのマルコフチェーンでは Accuracy が下がっていく結果となった。クラス毎の予測結果を F 値で見えていくと、最も高い値が得られたのは AP 番号 A で 0.464 となった。比較的高い値が得られたのは、AP 番号 N, C, D, F であった。これらの AP は表 3.1 に示す通り、サンプル数が多い AP であり、状態遷移確率が実際の猿の動きに近い値になっていると考えられる。農家に猿がどこに出現するか知らせるシステムとして考えた場合、Accuracy で 31.5 % では 3 回に 1 回の正解が出せる程度なので、いまの結果をそのままシステム化することは精度の問題で難しいが、現在の猿の出現地を基準にして次にどこに動くのかを確率で知らせる、天気の前確率のようなシステムであれば、十分に検討する価値のある精度とも考えられる。ただし、サンプル数が少ない AP については継続的な行動収集が必要であるが、出現日数が少ないということは被害に合う確率も少ないことに直結する。

2 重マルコフチェーンが最も良くなった理由は次のように考察できる。猿は 2, 3 日毎の周期での行動（移動や同じ AP への連続した出現）を繰り返しており、また、移動する際は移動する方向（辿っていく AP）がある程度決まっていると考えられる。また、連続した出現に適した AP や、あまり出現に適さない AP もある。このような傾向は図 3.3 から読み取ることができ、例えば AP 番号 F - N - D - C - B と北上していく経路があると想定できる。また、AP 番号 A や C は出現に適した餌場があると考えられ（実際の環境としても田畑が多い）、状態遷移先が同じ AP、つまり A - A のようなマルコフチェーンが多くなっている。連続した出現といっても、同じ AP に長期間に渡って居続けるという状態は、AP 番号 A において図 3.2 の 1 年目の 180 日前後に多く見られるが、基本的には稀な現象である。以上のような理由のため、2, 3 日毎の周期での行動を繰り返していると想定でき、本論で扱う猿の周期の特性があらわれやすい単純マルコフチェーンや 2 重マルコフチェーンにおいて予測結果が高くなる傾向があると考えられる。

表 3.4: マルチクラス問題の結果

Table 3.4 Result of multi-class problem

Evaluation	AP No.	1st	2nd	3rd	4th	5th
Precision	N	0.232	0.280	0.220	0.213	0.218
	A	0.318	0.382	0.300	0.333	0.333
	C	0.440	0.325	0.275	0.259	0.273
	D	0.283	0.308	0.308	0.286	0.375
	F	0.306	0.375	0.429	0.308	0.333
	B	0.000	0.500	0.000	0.000	0.000
	E	0.000	0.000	0.000	0.000	0.000
	Z	0.000	0.385	0.294	0.250	0.333
Recall	N	0.568	0.580	0.605	0.765	0.889
	A	0.318	0.591	0.273	0.136	0.045
	C	0.431	0.490	0.275	0.137	0.059
	D	0.289	0.267	0.178	0.089	0.067
	F	0.306	0.245	0.184	0.082	0.041
	B	0.000	0.036	0.000	0.000	0.000
	E	0.000	0.000	0.000	0.000	0.000
	Z	0.000	0.070	0.070	0.028	0.014
F-measure	N	0.330	0.378	0.322	0.333	0.350
	A	0.318	0.464	0.286	0.194	0.080
	C	0.436	0.391	0.275	0.179	0.097
	D	0.286	0.286	0.225	0.136	0.113
	F	0.306	0.296	0.257	0.129	0.073
	B	0.000	0.067	0.000	0.000	0.000
	E	0.000	0.000	0.000	0.000	0.000
	Z	0.000	0.119	0.114	0.051	0.027
Accuracy	-	0.282	0.315	0.249	0.225	0.225

3.6 結言

本章では猿の出現予測を行うために猿の動向を2年間に渡って収集した。実験システムを稼働させた2012年7月1日から2014年5月15日までの約2500万件の猿の信号を基に、集落を猿がどのように移動しているのかを1日毎のスパンで明らかにした。さらに実験期間を1年目、2年目に分け、1年目の移動データを教師として5次までの M 重マルコフチェーンの状態遷移確率を計算した。状態遷移確率を分類器として用いて、2年目の猿の移動データに対して、猿がどこのAPに出現するかの予測を行った。

猿が1日毎に集落を移動する様子をグラフとしてあらわしたところ、季節的な移動の特徴が見られた。夏から秋にかけては、同じ集落に連続して出現することは少なく活発に移動している。また、冬から春にかけては同じ集落に連続して出現することが多い。季節によって山や農作地に存在する食物量に違いがあり、気温が低い場合は移動のためのカロリーを節約するために長距離の移動が減るためと考えられる。

また、状態遷移図からは自己遷移の多い集落や移動の通り道になっていると考えられる集落の存在が見て取れ、猿には何らかの移動ルートが存在する可能性も示された。

実験の結果、2重マルコフチェーンの性能がもっとも良く、猿がどこかのAPに出現するかそれとも非出現になるかの2クラス問題では、Accuracyで目標値が80%に対して57.5%が得られた。猿が出現するかどうかは農家にとっては切実な問題であり、半分以上の割合で予測が当たるのであれば農家にとっても十分に有益な結果といえる。同様に、猿がどこのAPに出現するかのマルチクラス問題では、Accuracyの目標値が50%に対して31.5%となった。得られた精度では猿がどこのAPに出現するかを十分に当てられるとは言いがたいが、現在の猿の出現地を基準にして次にどこに動くのかを確率で知らせる、天気降水確率のようなシステムであれば、十分に検討する価値のある精度とも考えられる。

それぞれの目標値に対して結果が20%程度下回っているため、今後も改良を続ける必要がある。

第4章 環境要因を用いた猿の出現予測

4.1 緒言

前章では、マルコフチェーンモデルを用いて猿の移動を1日毎に観測することで猿が出現するかどうかは高い精度で予測でき、さらにどこの農作地に出現するかまでを予測できる可能性を示した。しかし、マルコフチェーンモデルを適用するためには、予測したい日の前日までの正確な猿の出現データを入力として与える必要がある。正確な猿の出現データを記録するという事は、対象とする農作地全てを人が監視し、猿が出現したかどうかを記録しなければならない。また、監視が疎かになった結果として非出現であったとする誤入力が続いてしまうと、実際には猿が出現するにも関わらず、非出現としか予測しないモデルになってしまう。

以上のジレンマを解消するためには、人が積極的に観測しなくても得られるパラメータを使い、猿の出現を予測できればよい。そこで本章では、気象データや曜日・時間などの環境パラメータを予測に使うことを検討する。環境パラメータは一般的かつ継続的に得られるデータであるため、発信機付きの首輪のように猿の捕獲や電池切れの心配をする必要がなくなり、また、猿の出現を逐次入力として与える必要もなくなる。

予測手法は、確率的手法であるベイズ推定と、教師データとのベクトル距離による判別手法であるSVMを用いる。ベイズ推定を用いる理由として、ベイズ推定を用いることで過去のデータをもとに未来の予測が可能であるため、猿の出現予測にも応用できる可能性がある。線形の問題であればベイズ推定で十分な予測精度になるが、非線形の問題であればベイズ推定では予測が成り立たない。そのため、非線形な分類が可能であるSVMを本論の主な予測手法として用いた。

4.2 原理

4.2.1 ベイズ推定による出現予測

ベイズ推定はベイズの定理を用いて、ある根拠に基づきその原因となった事象を推定する手法である。ベイズの定理は以下の式であらわされる。ここで X を予測に用いるパラメータ、猿出現の条件を A 、非出現の条件を N とすると猿が出現するときの条件付確率 $P(X | A)$ とは出現予測に用いる 8 種類の各パラメータ (4.4 節で後述する) の出現確率 $P(X_k | A)$ を教師データから求めることができる。猿が非出現であるときの条件付確率 $P(X | N)$ 、猿が出現する確率 $P(A)$ 、猿が非出現である確率 $P(N)$ も同様に教師データから求める。猿が出現するときの条件付確率 $P(X | A)$ は以下のようにして求める。まず、4.4 節で後述するように、パラメータを分割し、猿の出現時・非出現時でそれぞれ分割したパラメータ毎の出現率を求める。分割したパラメータ毎の出現率とは、離散値や、正規化された連続値を 8 分割したパラメータが対象期間中に何回観測されたかを求め、その回数を全体のスロット数で除算したものである。これを猿の出現時と非出現時でそれぞれ集計し、求める。求めた確率を使用してベイズ推定を行い、猿の出現する確率を求める。4.1 式により求めた $P(A | X)$ の値が一定以上だった場合は猿が出現すると判断し、そうでない場合は猿が出現しないと判断する。

$$P(A|X) = \frac{P(X|A)P(A)}{P(X)} \quad (4.1)$$

$$= \frac{P(X|A)P(A)}{P(X|A)P(A) + P(X|N)P(N)} \quad (4.2)$$

ここで、

$$P(X|A) = \prod_{k=1}^8 P(X_k|A) \quad (4.3)$$

$$P(X|N) = \prod_{k=1}^8 P(X_k|N) \quad (4.4)$$

4.2.2 SVM による出現予測

SVM[69] は教師あり学習の手法の一つである。SVM はクラス間のマージンを最大化することで最適な分離超平面を引くことができ、汎化能力が高いという特徴を持つ。また、SVM は線形分離できないデータだけでなく、非線形な分離を必要とするデータにも適応可能であるため、様々な分野で応用

されている [70][87][89]. 本研究では SVM を利用し, 猿の出現時と非出現時のデータを学習させ, 区別させることで猿の出現予測を行う. SVM にてデータセット D に属する未知パターンの特徴ベクトル \mathbf{x} を 2 クラスのいずれかに識別する識別関数を次式で表す.

$$f(\mathbf{x}) = \text{sign} \left(\sum_{i \in D} \alpha_i y_i K(\mathbf{x}_i, \mathbf{x}) + b \right) \quad (4.5)$$

ここで y_i は i 番目の学習サンプル x_i に対応するクラスラベルであり, α_i はラグランジュの未定定数, b はバイアス項である. $K(x_i, x)$ はカーネル関数であり, 次式であらわす Radial basis function (RBF) カーネルを用いる.

$$K(\mathbf{x}_i, \mathbf{x}) = \exp \left(-\gamma \|\mathbf{x}_i - \mathbf{x}\|^2 \right) \quad (4.6)$$

ここで γ はカーネルパラメータである. カーネル関数は線形分離不可能な学習データを, 高次元の特徴空間に写像し, 写像先の特徴空間において線形分離可能にする. そのため, 線形分離が難しいと考えられる猿の出現予測に適していると考えられる. 4.5 式に示した識別関数を求めるためには, マージン最大化基準によって次式を最大にする α_i を求める.

$$\max \sum_{i \in D} \alpha_i - \frac{1}{2} \sum_{i, j \in D} \alpha_i \alpha_j y_i y_j K(\mathbf{x}_i, \mathbf{x}_j) \quad (4.7)$$

$$\text{subject to } \sum_{i, j \in D} \alpha_i y_i = 0, 0 \leq \alpha_i \leq C \quad (4.8)$$

C はどの程度の誤識別を許すかを決定するペナルティパラメータ (コストパラメータ) であり, 学習時にあらかじめ求めておくパラメータはコストパラメータ C とカーネルパラメータ γ の 2 つである. 本研究では SVM のライブラリとして LIBSVM[82] を用いた. C と γ は $[2^{-15}, 2^{-14}, 2^{-13}, \dots, 2^{15}]$ と変化させていき, 最適な組み合わせを分類に用いた. SVM を用いて出現予測を行う手順は以下の通りである. まず, 4.4 節で後述するようにパラメータを正規化し, 猿の出現時と非出現時のパラメータをそれぞれ SVM に学習させる. 猿の出現を予測する際は, 予測する日時におけるパラメータを SVM に入力し, 出現と非出現のどちらのクラスに属するかを判定することで猿の出現を予測する. 本研究ではベイズ推定の結果と同じスケールで比較するために, LIBSVM のオプションを利用して 0~1 の実数値を出力し, この値が大きいほど猿の出現する確率が高いと考える.

4.2.3 複合型分類器

クラス分類の問題において、クラスのサンプル数がどちらかのクラスに多く偏り、アンバランスになることはよくある。このとき、アンバランスな数の教師データを SVM に学習させると、SVM は教師データ数の多いクラスに分類をしようとする。サンプル数を合わせる一般的な対処法としては、少ないクラスをオーバーサンプリングして増やしたり、多いクラスを少ないクラスに合わせてダウンサンプリングすることが挙げられる。本研究では、教師データのパターンを 2 種類用意し、組み合わせによって正解率を上げる新規手法を提案する。2 クラス間で分離が難しい領域があるサンプル群で、ポジティブなクラスが少ない、ネガティブなクラスが多いとする。このとき、図 4.1 のようにポジティブな教師データ数に合わせてネガティブのトレーニング数を決めると、ネガティブクラス全体の数に対して教師データ数が非常に少なくなるためにネガティブクラスの素性を捉えきれず、ポジティブクラスに優位な分類器になると考えられ、本論ではこれをポジティブバイアスと呼ぶことにする。逆に、ネガティブクラスの全体の数に合わせて教師データ数を決める際には、ポジティブのデータ数は簡単に増やすことができないため教師データの重みを増やすことで対応できるが、ポジティブクラスの素性が増えるわけではないため、2 クラス間で分離が難しい領域はネガティブに多く分類される結果となる。本論ではこれをネガティブバイアスと呼ぶことにする。そこで、2 つのバイアスの分類器を組み合わせることで正解率を上げる。

図 4.2 は 2 つのバイアスの分類器からどのようにクラスを識別するかを示している。本論では猿が出現することを高い精度で住民に知らせるために出現と判断するための基準を上げる。本章で取り扱う環境要因に対して猿の出現・非出現クラスは非線形な挙動を示すと考えられるため、2 クラス間で分離が難しい特徴領域を多く含んでいると考えられる。そのために複合型分類器が両方とも出現と判断した場合、つまり図 4.2 の右上領域の場合には出現クラスであると識別し、右上以外の領域では非出現クラスであると識別する。

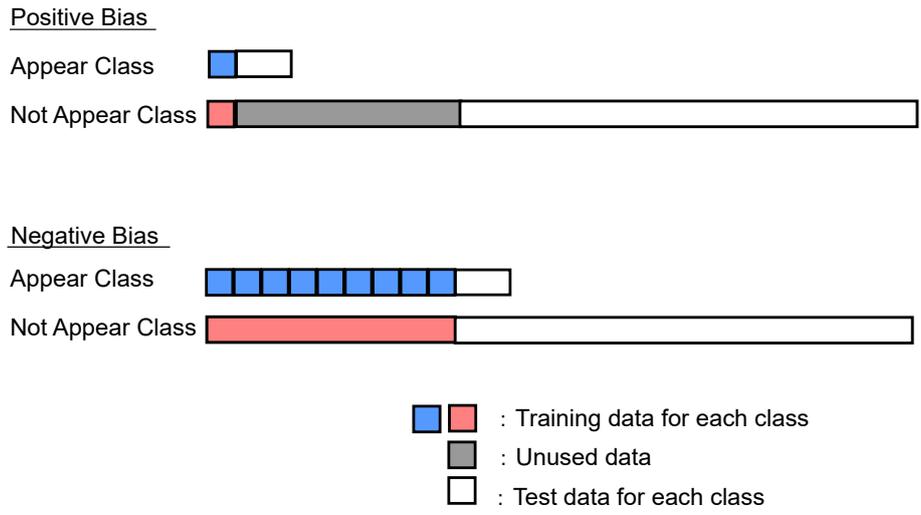


図 4.1: 複合型分類器のバイアス
 Fig. 4.1 Bias of combination classifier.

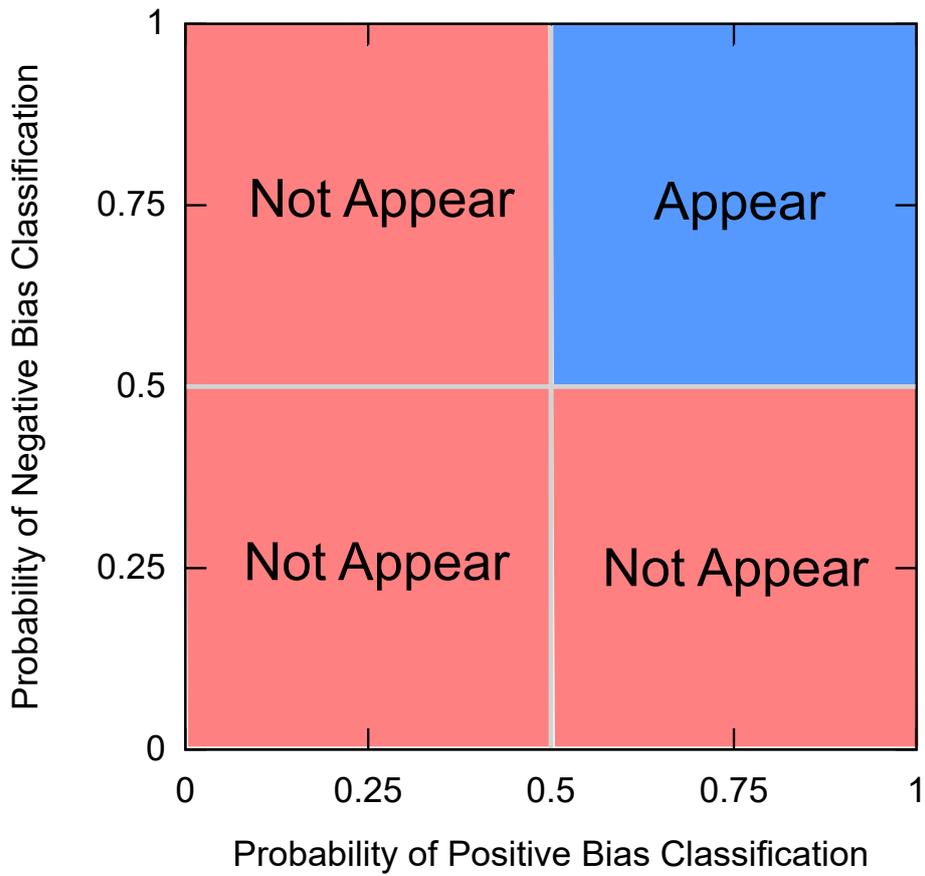


図 4.2: 複合型分類器の識別領域
 Fig. 4.2 Classification area of combination classifier.

4.3 猿の出現の定義の時間的分割

3.2.2 項では、猿がどの AP に出現するかというマルチクラス問題を取り扱うため、猿の出現を 1 日毎に定義した。本章では、猿の出現が何らかの環境パラメータに依存しているという考えのもと、環境パラメータの最小単位時間である 1 時間単位のスロットに猿の実験期間を分割し、スロット毎に猿の出現、非出現を判別する。スロットの時間に、猿が接近しているとみなせる RSSI のときに出現とカウントし、そうでない場合は非出現とカウントする。2012 年 7 月 1 日から 2014 年 5 月 15 日までを 1 時間単位のスロットに分割すると、16305 スロットとなった。これらのスロットにて出現、非出現をカウントしたところ、出現のスロットは 1714、非出現のスロットは 14591 となった。これら 16305 スロットのデータを猿の出現予測に使用する。

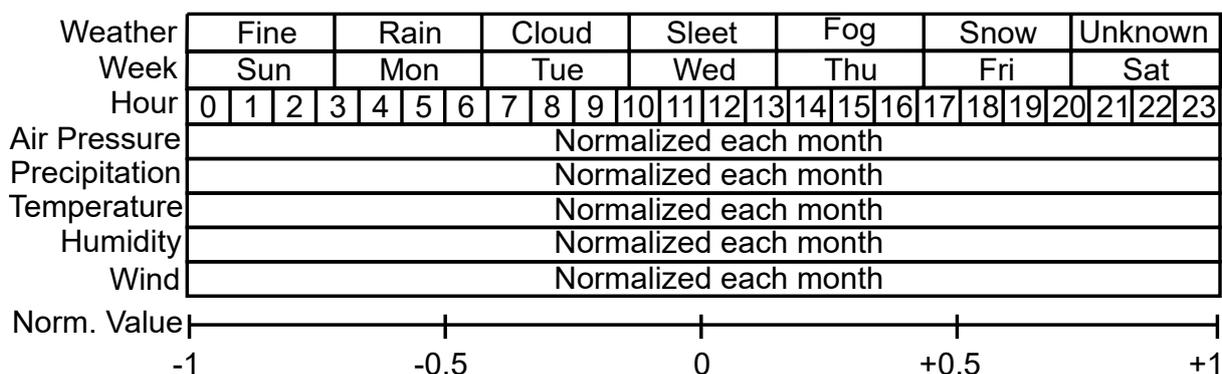


図 4.3: パラメータの正規化方法
Fig. 4.3 Normalization method for Parameters.

4.4 予測で使用するパラメータの分析

出現予測に使用するパラメータとして、気象データ(天気、湿度、降雨量、気圧、気温、風速)と時刻と曜日を使用する。気象データは気象庁提供のものを用いた。Unknown は特定の気象条件が示されていないことを意味する。パラメータを図 4.3 に示す。猿が活動しやすい時間帯や、夜など活動しない時間帯がある場合に対応できるように時刻をパラメータとして導入した。また、休日は平日よりも人が外にいる割合が高く、猿が農作地へ近づきにくくなると考えられるため曜日をパラメータとして導入した。

対象期間が長いため、気象データは一ヶ月毎に最大値と最小値を求めて-1 ~ +1 に正規化した。ベイズ推定ではパラメータ毎に出現、非出現の尤度を求める必要がある。そこで、離散値である天気、時刻、曜日はそのままの分割数を使い、気温や湿度のような数値毎に出現回数を求めることが難しいものは正規化したものを8つに分割した。事前実験として分割数2, 4, 8, 16, 32 と試した結果から最も良い結果が得られた8分割としている。また、SVM に離散値を入力する場合は、-1 ~ +1 に収まるように分割数毎に値をふった(例えば0時は-1.0, 1時は-0.913, 23時は+1.0のようにした)。

図 4.4 は猿の実験データ 16305 スロットにおけるパラメータの出現/非出現の比をあらわしたものである。縦軸は、それぞれのパラメータにおける出現/非出現の比をあらわす。連続値は本節で説明しているように、8分割している。各パラメータの傾向を分析していく。天気において、雨、みぞれ、霧のときに出現率が高いのが目立つ。3項で述べたように、猿は視界が悪い状況を好んで農場に出現する可能性があり、結果にあらわれていると考えられる。同様に、降雨量が多くなるほど出現率があがっていることも仮説を支持している。時間については夕方から夜にかけての暗い時間に出現の割合が増え、昼間は非出現の割合が増えている。温度についての相関としてよくあらわれていると考

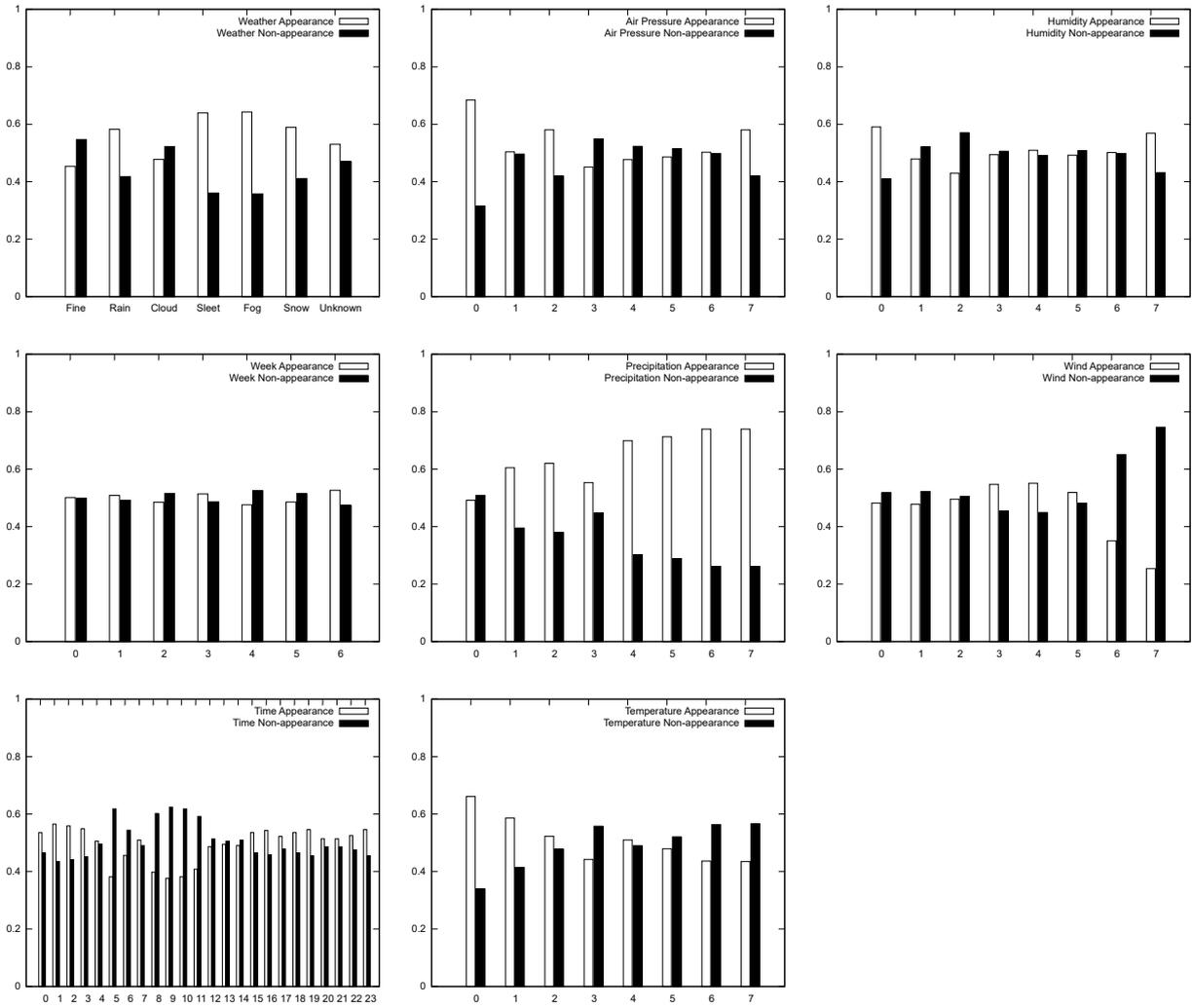


図 4.4: 猿の実験データ 16305 スロットにおけるパラメータの出現/非出現の比

Fig. 4.4 The parameter ratios of the appearance and the non-appearance of monkeys in investigated 16305 time slots.

えられる（つまり、昼間ほど温度が高い）気圧では、中央値から離れるほど出現が増えているが、穏やかでない天気ほど猿にとっては出現しやすいと考えられる。風速については、強い風速を境に非出現の割合が急に増えている。湿度については定性的な考察が難しいが、湿度が上がる条件は温度が低い、雨が降っていることが挙げられるので、湿度が高いときに猿の出現が増えるのはこれらの条件に当てはまると言える。曜日については残念ながら特徴的な相関は見られない。以上のようにパラメータ毎の相関を考察したが、猿は生物である以上、特定の環境条件のときに必ず出現するとは言えず、非常に非線形な挙動をすることが考えられる。

4.5 実験方法

4.3 項で述べたように，16305 件の時間スロットデータを準備した．スロットデータのうち，1714 件がポジティブクラスである「出現」データ，14591 件がネガティブクラスである「非出現」データである．今回のサンプル群から 600 件ずつをポジティブバイアス試験の教師データとして使う．ネガティブバイアス試験の教師データには，ネガティブクラスからはポジティブバイアス試験と同じ 600 件を含む 5400 件を選び，ポジティブクラスからはポジティブバイアス試験と同じ 600 件を使い，トレーニングの重みを 9 倍することで対応する．分類にかけるサンプルはポジティブクラスでは 1114 件，ネガティブクラスでは 9191 件となる．ベイズ推定，SVM とともに 0.5 以上の確率の場合に出現とし，それ以下の確率では非出現とした．評価は正解率で行うが，4.9 式で計算する．

$$\text{正解率} = \frac{TP}{TP + FN}, \frac{TN}{TN + FP} \quad (4.9)$$

表 4.1: ベイズ推定と SVM の正解率

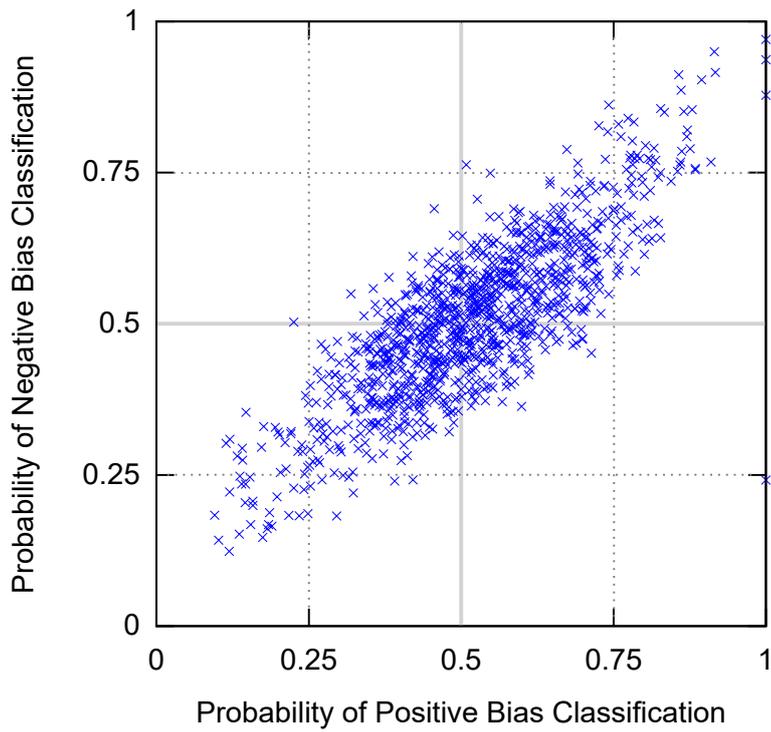
Table 4.1 Precision Result of the Bayesian estimation and Support Vector Machine.

Method	Appearance (%)	Non-appearance (%)
Bayes Pos. Bias	12.5	90.7
Bayes Pos. + Neg. Bias	14.0	90.7
SVM Pos. Bias	20.5	92.4
SVM Pos. + Neg. Bias	31.0	91.7

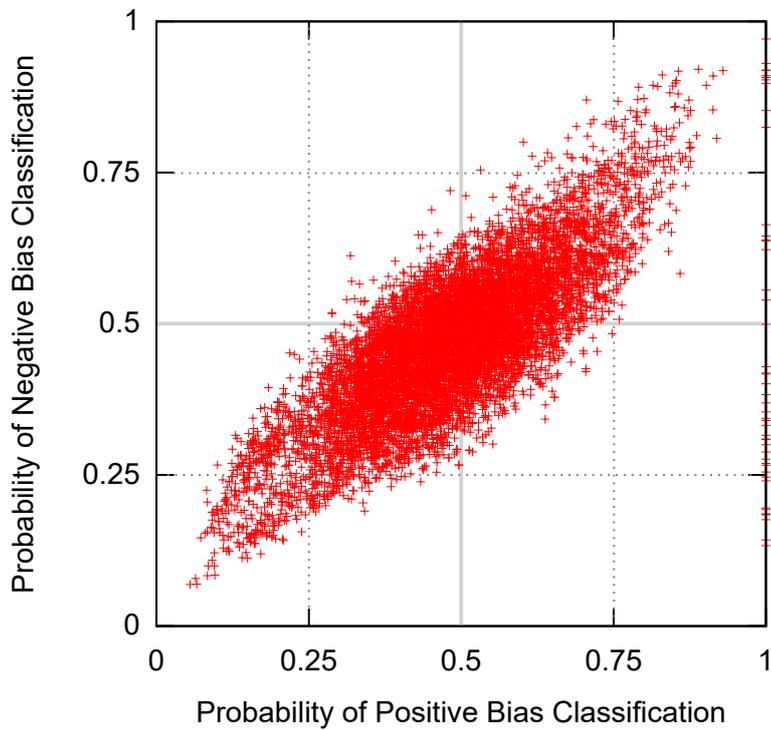
4.6 実験結果

ベイズ推定の実験結果を図 4.5 と表 4.1 に示す。図 4.5 は横軸にポジティブバイアス試験で分類したときの確率，縦軸にネガティブバイアス試験で分類したときの確率をとってプロットした。(a) は出現サンプルを分類した結果，(b) は非出現サンプルを分類した結果である。図 4.5 に示す通り，ベイズ推定では確率 0.5 を中心にガウス分布となってしまう，ほとんどクラス分類ができていないことがわかる。そのため，出現と分類した正解率はポジティブバイアス試験のみで 12.5%，ポジティブバイアス試験とネガティブバイアス試験の組み合わせでも 14.0% と非常に低い数値となった。図 4.3 では相関があるように見えても，猿の行動は予想通り非線形であると考えられ，ベイズ推定は出現予測に適していないことがわかった。

SVM の実験結果を図 4.6 と表 4.1 に示す。パラメータ γ と C との最適な組み合わせはポジティブバイアス試験のときに (16, 1)，ネガティブバイアス試験のときに (4, 0.5) であった。図 4.6 において，ポジティブバイアス試験での分類結果である横軸では 0 や 1 に近いような強い分類結果は出ないが，これは教師データが少ないことから予想どおりである。また，ネガティブバイアス試験での分類結果である縦軸では 0 や 1 に近い結果が出ており，(a) においても 0 に近いサンプルが多いことから，非出現クラスの教師データの増加に伴って，非出現と分類されるサンプルが増えることがわかる。しかし，出現と分類した正解率はポジティブバイアス試験のみの 20.5% からポジティブバイアス試験とネガティブバイアス試験の組み合わせでは 31.0% に増加しており，本研究の手法が有効であることがわかった。また，ベイズ推定のポジティブバイアス試験とネガティブバイアス試験の組み合わせでは出現の検出率は 41.6% であったが，ベイズ推定がガウス分布になったので順当な値といえる。ベイズ推定に比べて SVM のポジティブバイアス試験とネガティブバイアス試験の組み合わせでは出現の検出率は 31.3% となり，非出現の教師データに 9 倍の偏りがあるにもかかわらず，良好な結果が得られたといえる。

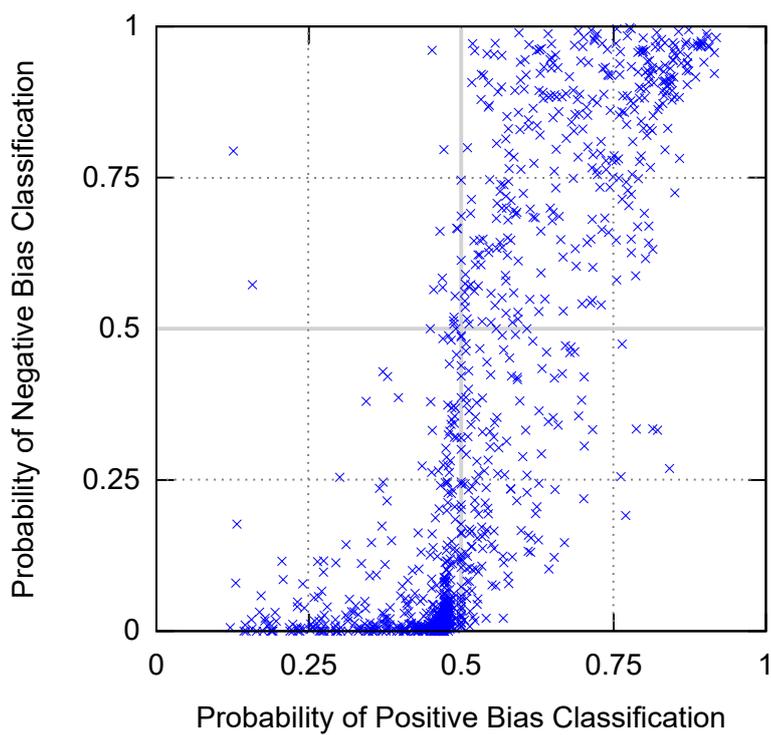


(a)

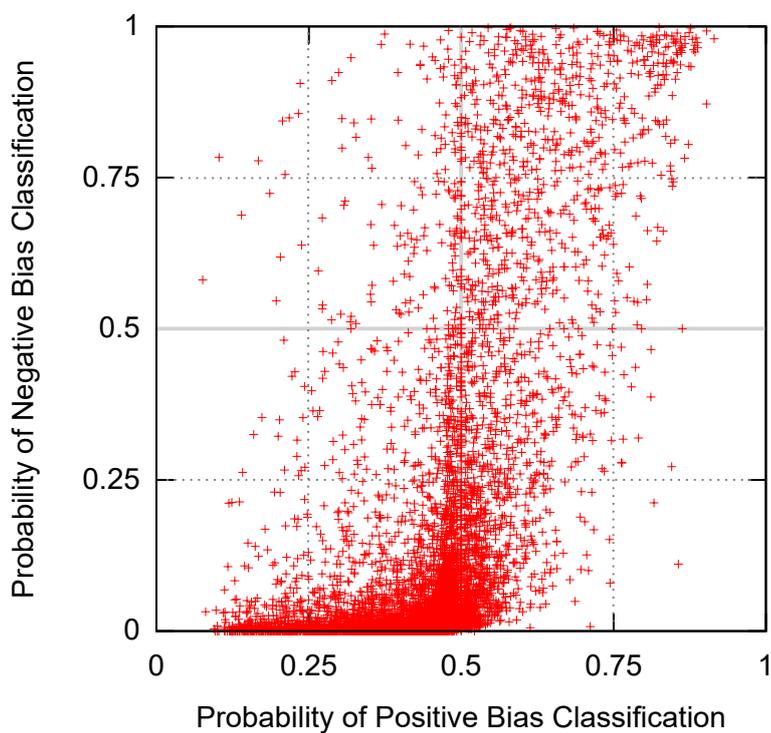


(b)

図 4.5: ベイズ推定の確率分布 (a) は出現サンプルの結果 (b) は非出現サンプルの結果
 Fig. 4.5 Probability graph of Bayesian estimation. (a) is the result of the appearance samples, (b) is the result of the non-appearance samples.



(a)



(b)

図 4.6: SVM による推定の確率分布 (a) は出現サンプルの結果 (b) は非出現サンプルの結果
 Fig. 4.6 Probability graph of Support Vector Machine. (a) is the result of the appearance samples, (b) is the result of the non-appearance samples.

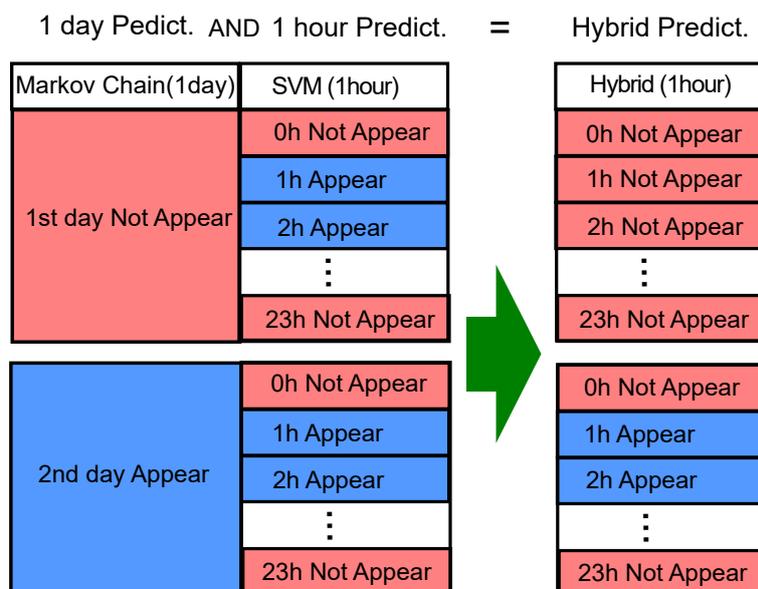


図 4.7: 移動・環境要因の組み合わせによる予測

Fig. 4.7 Hybrid Prediction of Transitions and Environmental Conditions.

4.7 移動要因と環境要因を組み合わせた出現予測

4.7.1 移動要因と環境要因の組み合わせ予測手法

3章では移動要因を用いて猿の1日毎の出現予測を行った。4章では環境要因を予測のファクターとして用いて、1時間毎にどこかのAPに出現するかの予測を行った。2つの手法での予測結果はそれぞれの要因に基づく独立した予測である。そのため、2つの手法を組み合わせることで出現予測の信頼性を上げられる可能性がある。4.7節では、移動・環境要因を組み合わせた場合の出現予測について述べる。

まず、図4.7のように、移動要因による1日単位の予測と、環境要因による1時間単位の予測を同じ時系列に揃える。これらは、同じ観測期間のデータを使用しているため、時系列を揃えるために特別な操作は必要ない。次に、それぞれの予測のANDをとり、移動要因と環境要因の組み合わせの結果として、1時間単位の出現予測が得られる。

4.7.2 移動要因と環境要因の組み合わせ予測結果

移動要因の予測結果として、3.4節の2重マルコフチェーンの結果を用いた。2重マルコフチェーンの出現予測の正解率は57.5%である。環境要因の予測結果として、4.6節のSVMの結果を用いた。ただし、2重マルコフチェーンの予測は猿の実験期間の2年間のうち、後半の2年目を対象としているため、SVMの結果も2年目のみを抽出して用いた。以上により、テストデータとしては365日で

表 4.2: 移動要因と環境要因の組み合わせによる正解率

Table 4.2 Precision Result of the Combination of Transtions and Environmental Conditions.

Method	Appearance (%)	Non-appearance (%)
SVM only	34.4	91.1
Markov Chain + SVM	39.8	90.5

あり、1時間単位のタイムスロットとしては出現が647スロット、非出現が4927スロットであった。4.7.1項で説明したようにマルコフチェーンとSVMの予測結果のANDをとることで組み合わせの予測結果とし、4.9式に従って正解率を算出した。

表 4.2 に正解率を示す。実験期間の2年目のみを抽出した場合、SVMによる出現の正解率は34.4%であった。マルコフチェーンとSVMのANDをとった結果、出現の正解率は39.8%となり、SVM単独に比べて5.4%の上昇となった。このことから、移動要因と環境要因の組み合わせは出現予測に効果的であると考えられる。

4.8 結言

対象期間を1時間毎のスロットに分けると16305スロットとなり、スロットにおいて猿が出現したかどうか、RSSIを基準にラベル付けを行った。ラベル付けされたスロットに時刻、曜日、気象を合わせ、ベイズ推定とSVMを用いて猿の出現予測を行った。また、正解率を向上させるため、2つの学習パターンを組み合わせる新規手法を提案した。本研究の目的である猿の出現予測を行った結果、ベイズ推定では出現の正解率は14.0%にとどまったが、SVMでは本研究の手法において31.0%の正解率が得られた。実験結果では、非常に非線形な挙動を示すと考えられる猿の行動を、比較的簡単に得られる環境条件をパラメータにした結果としては十分だと考えられる。また、現在までは農作物被害を予防するための猿の出現予測の手法が確立されていなかったが、本研究の結果、直接的な指標がなくても猿の出現を約3割の確率で当てられる可能性が示された。

また、移動要因と環境要因はそれぞれ独立した要因として本論で検証してきたが、これらの要因を組み合わせることで猿の出現を約4割の確率で当てられる可能性も同様に示された。

第5章 結論

5.1 本論のまとめ

本論では、猿の出現を予測するシステムを作成することを目的とした。

予測に先立って、猿の行動を記録する実験システムを構築し、三重県伊賀市において2年間に渡って稼働させた。実験システムの稼働率は高い水準であり、猿の行動を多種のデータとして取得できた。実験データをもとに猿が農作地に出現したかどうかのラベル付けを行い、マルコフチェーンモデルによる1日毎の出現予測と、SVMによる1時間毎の出現予測を行った。

マルコフチェーンの状態遷移確率を計算し、単純マルコフチェーンを状態遷移図として表現したところ、集落間の距離や集落毎の特徴により猿の移動も特徴付けられる可能性が示された。ひとつの特徴として、同じ集落に連続して出現することが確認された。つまり自己遷移が多い集落は、猿にとって有益な餌場になっている可能性が高く、理由としては農作地の作付け量が豊富であったり、集落内での人による追い払いが少ないために猿にとって採食のリスクが低いことが考えられる。また、猿の進行ルートもいくつか想定できることから、採食行動に向く集落や移動の通り道だけの集落など、猿にとって集落の利用方法が違う可能性も示された。

マルコフチェーンによる予測では、単純マルコフチェーンから5重マルコフチェーンまでを使って予測精度を比較した。実験結果では、2重マルコフチェーンの性能がもっとも良く、猿がどこかのAPに出現するかそれとも非出現になるかの2クラス問題では、Accuracyで目標値が80%に対して57.5%が得られた。猿が出現するかどうかは農家にとっては切実な問題であり、半分以上の割合で予測が当たるのであれば農家にとっても十分に有益な結果といえる。同様に、猿がどこのAPに出現するかのマルチクラス問題では、Accuracyの目標値が50%に対して31.5%となった。AP毎のAccuracyを比較したとき、マルコフチェーンの状態遷移確率を計算するための教師データ数が多いAPほど高いAccuracyが得られた。教師データ数が多いほど、実際の猿の移動を近似したモデルになっていると考えられる。現在の精度では猿がどこのAPに出現するかを十分に当てられるとはいえず、現在の猿の出現地を基準にして次にどこに動くのかを確率で知らせる、天気降水確率のようなシステムであれば、十分に検討する価値のある精度とも考えられる。

本論では環境要因の分析結果を基に、非線形学習器のひとつである SVM を用いて 1 時間毎の出現予測を行った。SVM では本論の手法において 31.0 % の正解率が得られた。実験結果として、非常に非線形な挙動を示すと考えられる猿の行動を、比較的簡単に得られる環境条件をパラメータにした結果としては十分だと考えられる。また、現在までは農作物被害を予防するための猿の出現予測の手法が確立されていなかったが、本研究の結果、直接的な指標がなくても猿の出現を約 3 割の確率で当てられる可能性が示された。

環境要因による出現予測を行うにあたり、本論で用いる環境要因と猿の出現・非出現の割合の比較を行った。本論で取り扱う環境要因の多くは単独の要因では猿の出現と線形的な関係を持ちづらく、非線形的な挙動になっている。猿の行動調査に関する先行研究で得られていた知見として、猿は人間との遭遇を避ける傾向が指摘されている。本論の実験結果でも、降雨時や夜間時など、人間が屋外で活動しづらい環境のときに猿が出現しやすい傾向が得られた。

移動要因と環境要因はそれぞれ独立して精度を検討できる予測要因・手法ではあるが、2 つの手法を組み合わせると出現予測を行った結果、予測の精度を向上させることができた。移動要因を用いた予測手法で猿の現在位置を絞り込んでいくが、何時のタイミングで集落に出現するかどうかの予測までは難しい。そこで、環境要因を用いた予測手法で集落に出現するかどうかの判定を 1 時間単位で予測するという方法を組み合わせることで、獣害対策のための出現予測手法として有効な手法を提案できた。

5.2 今後の展望

本論では、移動を要因に用いた予測としてマルコフチェインモデルを使って猿の移動をモデル化することを試みたが、猿の移動結果のみを特性として取り上げており、なぜ移動したのかという理由にまでは言及できていない。なぜ猿が別の場所に移動したのか、それとも同じ場所に出現したのかという隠れマルコフ状態は取り扱っていない。そこで今後の課題として、なぜ移動したのかという隠れマルコフ状態を推定する、もしくは現地実験をさらに重ねて実際の移動に関するファクターを特定することができれば、さらに予測の精度を上げられる可能性がある。

さらには、本論で対象とした地域には移動経路になっているだけの場所も存在する可能性が考えられるため、出現予測の対象地域とするかどうかは今後の議論が必要となる。また、遷移の確率を左右するパラメータとして、農作地周辺の作物の収量を季節毎に組み込むこともひとつの手段として考えられる。農作地だけでなく、季節によって山に存在する木の実や新芽など猿にとっての餌の生育量が変わるため、それによっても農作地に出現する割合も変わる可能性が高い。そのため、季節毎のマルコフチェインモデルを用意することができれば、さらに精度向上に貢献できると考えられる。

環境パラメータを要因に用いた予測については、気象庁から得られた気象データを使用しているため、対象地域の気象条件を正確に反映できていない可能性もある。気象庁の発表する気象データは、その地域を代表する都市の観測データであるが、本論が対象とする猿の出現地域は気象データと同一の市であるが山間部の農作地であり、都市部に比べれば気象変動が大きいためである。そこで、気象庁の気象データと猿の出現地域の気象データとの相関を検証することで、予測精度を高められると考えられる。

また、本論では猿の群れ1つについての実験結果をもとにした予測を行っている。ある地域に複数の群れが存在する状況においては群れの行動パターンが変わってしまうことも考えられるため、継続的な行動の収集が必要である。猿の群れの移動を長期的に観測し続けるのは実験機器の電池寿命の問題から難しいため、出現リスクの高いAPに予測範囲を絞ることで、予測の精度を高めることも検討の余地がある。ただし、本論は獣害対策のために出現予測を行っているという性質上、システム側の都合で対象地域を絞ってしまうのは得策ではなく、住民とも相談の上で対象地域を決定する必要がある。

謝辞

本研究を進めるにあたり，終始有益なご指導とご助言を下された静岡大学創造科学技術大学院情報科学専攻教授杉浦彰彦先生に深く心から感謝いたします。

また，本論を書く上で貴重なお時間を頂戴し，様々な助言をくださった静岡大学創造科学技術大学院情報科学専攻教授中谷広正先生，静岡大学創造科学技術大学院情報科学専攻教授佐治斉先生，静岡大学創造科学技術大学院ナノビジョン工学専攻講師武田正典先生に感謝の意を表したいと思います。

また，学位取得に協力していただいた鳥羽商船高等専門学校制御情報工学科学科長江崎修央先生をはじめ，鳥羽商船高等専門学校の先生方に深く感謝いたします。

また，研究に協力してくださった杉浦研究室，江崎研究室，中井研究室の皆様にも心からの感謝を捧げます。

参考文献

- [1] 山端直人, “集落ぐるみのサル追い払いによる農作物被害軽減効果-三重県内 6 地区での検証-,” 農村計画学会誌, pp.273–278, feb 2010.
- [2] 中川尚史, “金華山のニホンザルの定量的食物品目リスト : 付記 : 霊長類の食性調査法と記載法の傾向,” 霊長類研究, vol.13, no.1, pp.73–89, may 1997.
- [3] T. Yamato, F. Shiho, S. Hideki, S. Chiemi, and T. Seiki, “Long-term variation in fruiting and the food habits of wild japanese nacaques on kinkazan island, northern japan,” American Journal of Primatology, vol.68, no.11, pp.1068–1080, 2006.
- [4] 深谷もえ, 岩本俊孝, 伊藤 哲, “幸島の植物種および植生タイプに対するニホンザルの採食選択性,” 霊長類研究, vol.18, no.2, pp.203–213, oct 2002.
- [5] 江成広斗, 松野葉月, 丸山直樹, “白神山地北東部に生息する野生ニホンザル (*macaca fuscata*) の農地利用型食物選択,” 野生生物保護, vol.9, no.2, pp.77–92, jan 2005.
- [6] 吉田 洋, 林 進, 北原正彦, 藤園 藍, “富士北麓地域におけるニホンザル野生群による農作物被害と被害防除の実態,” 農村計画学会誌, vol.25, no.2, pp.111–119, sep 2006.
- [7] 辻 大和, “ニホンザルの食性の種内変異 : 研究の現状と課題,” 霊長類研究, vol.28, no.2, pp.109–126, dec 2012.
- [8] 大井 徹, “採食行動の比較からツキノワグマとニホンザルの種間関係を考える,” 霊長類研究, vol.29, no.2, pp.123–135, 2013.
- [9] 大井 徹, 中下留美子, 藤田昌弘, 菅井強司, 藤井 猛, “西中国山地のツキノワグマの食性の特徴について,” 哺乳類科学, vol.52, no.1, pp.1–13, jun 2012.
- [10] 齊藤 v 千映美, 森光由樹, 清野紘典, “実験的環境変化がニホンザル (*macaca fuscata*) の行動圏利用に与える影響,” 哺乳類科学, vol.46, no.1, pp.63–64, jun 2006.

- [11] 山端直人, “集落ぐるみの追い払いがサル群の行動域や出没に与える効果,” 農村計画学会誌, vol.30, pp.381–386, nov 2011.
- [12] H. Enari and T. Suzuki, “Risk of agricultural and property damage associated with the recovery of Japanese monkey populations,” *Landscape and Urban Planning*, vol.97, no.2, pp.83–91, 2010.
- [13] 永田幸志, “丹沢山地札掛地区におけるニホンジカの行動圏特性,” 哺乳類科学, vol.45, no.1, pp.25–33, jun 2005.
- [14] 石塚 讓, 川井裕史, 大谷新太郎, 石井 亘, 山本隆彦, 八丈幸太郎, 片山敦司, 松下美郎, “季節, 時刻および植生が大阪のニホンジカ (*cervus nippon*) の行動圏に及ぼす影響,” 哺乳類科学, vol.47, no.1, pp.1–9, jun 2007.
- [15] 江成広斗, “東日本におけるニホンザルの分布変化に影響する社会・環境要因,” 哺乳類科学, vol.53, no.1, pp.123–130, 2013.
- [16] 有本 勲, 岡村 寛, 小池伸介, 晃司, 梶 光一, “集落周辺に生息するツキノワグマの行動と利用環境,” 哺乳類科学, vol.54, no.1, pp.19–31, 2014.
- [17] 山端直人, “獣害対策の進展が農家の農地管理意識に及ぼす効果,” 農村計画学会誌, vol.29, pp.245–250, 2010.
- [18] 中村大輔, 吉田 洋, 松本康夫, “野生サル加害群の頭数に対する住民の意識構造,” 農村計画学会誌, vol.33, pp.287–292, 2014.
- [19] 山田 彩, “追い払い手段の異なる集落間におけるニホンザルの逃走開始距離の差異,” 霊長類研究, vol.28, no.1, p.13, 2012.
- [20] 山口 薫, 山路永司, “野生サル追い払い犬事業の成果と課題,” 農村計画学会誌, vol.33, pp.281–286, 2014.
- [21] 田中俊明, 室山泰之, “超音波音およびシチメンチョウ音声に対するニホンザルの忌避反応,” 霊長類研究, vol.31, no.1, pp.31–37, 2015.
- [22] 宇野裕之, 玉田克巳, 平川浩文, 赤松里香, “Gps テレメトリーの測位成功率及び測位精度の評価,” 哺乳類科学, vol.42, no.2, pp.129–137, dec 2002.

- [23] 佐伯 緑, 早稲田宏一, “ラジオテレメトリを用いた個体追跡技術とデータ解析法,” 哺乳類科学, vol.46, no.2, pp.193–210, dec 2006.
- [24] 佐伯 緑, 斎藤昌幸, 園田陽一, “個体追跡調査における携帯 gps 端末を使った位置情報システムの利用,” 哺乳類科学, vol.49, no.2, pp.277–282, 2009.
- [25] 本間謙也, 間所洋和, 佐藤和人, “動線解析によるイベント会場での行動パターン分類(パターン認識),” 電子情報通信学会論文誌. D, 情報・システム, vol.95, no.10, pp.1848–1858, oct 2012.
- [26] L. Zhang and Q. Ji, “Image segmentation with a unified graphical model,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.32, no.8, pp.1406–1425, Aug. 2010.
- [27] 池田思朗, “Hmm の構造探索による音素モデルの生成,” 電子情報通信学会論文誌. D-II, 情報・システム, II-情報処理, vol.78, no.1, pp.10–18, jan 1995.
- [28] 青木茂樹, 大西正輝, 小島篤博, 福永邦雄, “Hmm による行動パターンの認識,” 電子情報通信学会論文誌. D-II, 情報・システム, II-パターン処理, vol.85, no.7, pp.1265–1270, jul 2002.
- [29] 川口和英, 田口 想, 渡辺仁史, “ハイブリッド水族館内における人間行動分析に関する研究: 集客施設内の情報携帯端末利用による行動解析,” 日本建築学会計画系論文集, no.573, pp.49–56, nov 2003.
- [30] 斎藤参郎, 山城興介, 中嶋貴昭, “隠れマルコフモデルによる地区間歩行者流動パターンの動的推定-福岡市都心部での大名・国体道路・今泉地区への適用-,” 地域学研究, vol.34, no.3, pp.69–90, 2003.
- [31] 浅原彰規, 丸山貴志子, 佐藤暁子, “混合マルコフモデルに基づく歩行者動線解析方式,” 情報処理学会論文誌, vol.52, no.1, pp.187–196, jan 2011.
- [32] 浅原彰規, 佐藤暁子, 丸山貴志子, “混合自己回帰隠れマルコフモデルによる歩行者行き先予測,” 情報処理学会論文誌, vol.53, no.1, pp.342–351, jan 2012.
- [33] 岡本正紀, 吉田 裕, “移動体通信網トラヒック特性評価のためのシミュレーション解析法,” 電子情報通信学会論文誌. B, 通信, vol.84, no.4, pp.750–756, apr 2001.
- [34] 石井貴幸, 吉田 裕, “マルコフ移動モデルの移動ゾーン併合に関する検討,” 電子情報通信学会論文誌. B, 通信, vol.89, no.2, pp.299–302, feb 2006.

- [35] 石井貴幸, 吉田 裕, “二重マルコフ連鎖に基づくマルコフ移動モデルの拡張,” 電子情報通信学会論文誌. B, 通信, vol.90, no.7, pp.705–707, jul 2007.
- [36] 兼田敏之, 吉田琢美, “歩行者回遊行動のエージェントモデリング,” オペレーションズ・リサーチ: 経営の科学, vol.53, no.12, pp.672–677, dec 2008.
- [37] 西野 至, 西井和夫, 北村隆一, “観光周遊行動を対象とした複数目的地の組合せ決定に関する逐次的モデル,” 土木計画学研究・論文集, vol.17, pp.575–581, 2000.
- [38] 笠原秀一, 田村和範, 飯山将晃, 棕木雅之, 美濃導彦, “行動履歴に基づく地域の環境要因を考慮した観光行動モデルの構築とその応用,” 情報処理学会論文誌, vol.57, no.5, pp.1411–1420, may 2016.
- [39] A. Daniel and S. Thad, “Using gps to learn significant locations and predict movement across multiple users,” Personal Ubiquitous Comput., vol.7, no.5, pp.275–286, Oct. 2003.
- [40] 山崎俊彦, アンドリューギャラガー, ツーハンチェン, 相澤清晴, “季節・時刻を考慮した大規模位置履歴からの都市内旅行予測・推薦,” 電子情報通信学会論文誌 D, vol.97-D, no.9, pp.1437–1444, 2014.
- [41] 丹後俊郎, 阿部一洋, 狩野紀昭, “個人差を導入したマルコフモデル,” 応用統計学, vol.19, no.1, pp.1–18, 1990.
- [42] 小澤暁人, 吉田好邦, “マルコフ連鎖を用いた生活行動再現による家庭エネルギー需要の推定,” 環境情報科学論文集, vol.27, pp.97–102, 2013.
- [43] 高山純一, 杉山智美, “吸収マルコフ連鎖を用いた観測交通量からの od 推計法に関する研究,” 土木学会論文集, vol.569, pp.75–84, jul 1997.
- [44] 岡村寛之, “マルコフ連鎖の極限推移確率と web リンク解析,” オペレーションズ・リサーチ: 経営の科学, vol.54, no.12, pp.739–743, dec 2009.
- [45] L. Page, S. Brin, R. Motwani, and T. Winograd, “The pagerank citation ranking: Bringing order to the web.,” Technical Report 1999-66, Stanford InfoLab, Nov. 1999. Previous number = SIDL-WP-1999-0120.

- [46] 原口一馬, 島田伸敬, 白井良明, 三浦 純, “センサの見通し可能性を考慮した確率的モデルによる障害物地図の作成,” 日本ロボット学会誌, vol.25, no.4, pp.632–642, 2007.
- [47] 小林 太, 安田裕二, 小島史男, “ベイズ推定を用いた自律搬送システム,” システム制御情報学会論文誌, vol.16, no.10, pp.500–506, 2003.
- [48] 大園 渉, 室町泰徳, “ベイズ推定法を用いた買物交通手段選択モデルの構築に関する研究,” 土木計画学研究・論文集, vol.25, pp.717–722, 2008.
- [49] K. Nobuo and I. Seigo, “Bayesian based location estimation system using wireless lan,” Third IEEE International Conference on Pervasive Computing and Communications Workshops, PerCom 2005 Workshops, 8-12 March 2005, Kauai Island, Hawaii, pp.273–278, mar 2005.
- [50] 伊藤誠悟, 河口信夫, “アクセスポイントの選択を考慮したベイズ推定による無線 lan ハイブリット位置推定手法とその応用,” 電気学会論文誌C (電子・情報・システム部門誌), vol.126, no.10, pp.1212–1220, 2006.
- [51] 松月正樹, 竹田 馨, 細川裕貴, 野呂雄一, 竹尾 隆, “無線 lan を用いた医療機器管理用 rfid の位置推定,” 電気学会論文誌C (電子・情報・システム部門誌), vol.133, no.12, pp.2143–2148, 2013.
- [52] 高橋誠也, 太田正哉, 佐々木純, 山下勝己, “ベイズ推定に基づく屋内避難誘導のための wifi 位置推定法,” 電気学会論文誌C (電子・情報・システム部門誌), vol.136, no.11, pp.1553–1554, 2016.
- [53] 加治 充, 原 隆浩, 西尾章治郎, “無線 lan による位置推定のためのオンライン生成可能な電波環境地図とその特性,” 情報処理学会論文誌, vol.50, no.1, pp.51–63, jan 2009.
- [54] 佐藤弘和, 伊藤誠悟, 河口信夫, “指向性アンテナを用いた無線 lan アクセスポイントの位置推定手法,” 電子情報通信学会論文誌, vol.88, no.11, pp.1634–1647, nov 2005.
- [55] 本村陽一, “ベイジアンネットワークによる日常生活行動モデリング,” 電子情報通信学会誌, vol.93, no.9, pp.774–778, sep 2010.
- [56] 手塚繁樹, 相馬 仁, 谷藤克也, “ベイジアンネットワークを用いた車線変更時の運転行動推定モデルの研究,” 日本機械学会論文集 C 編, vol.73, no.730, pp.1794–1800, 2007.

- [57] 周 洪鈞, 坂根茂幸, “ベイジアンネットワークの構造学習と推論を用いた移動ロボット位置決めのためのセンサプランニング,” 日本ロボット学会誌, vol.22, no.2, pp.245–255, 2004.
- [58] 石田千枝, 島田英一, 荒川 豊, 竹森敬祐, 笹瀬 巖, “ベイズ推測を用いた不正侵入イベント増減予測,” 情報処理学会論文誌, vol.46, no.11, pp.2704–2713, nov 2005.
- [59] 植野真臣, “e ラーニングにおける所要時間データの異常値オンライン検出 (教育工学),” 電子情報通信学会論文誌. D, 情報・システム, vol.90, no.1, pp.40–51, jan 2007.
- [60] 田中佑典, 上田修功, 田中利幸, “クラス固有の特徴選択に基づくベイズ識別器 (パターン認識),” 電子情報通信学会論文誌. D, 情報・システム, vol.96, no.11, pp.2755–2764, nov 2013.
- [61] 岩村雅一, 大町真一郎, 阿曾弘具, “パターン認識における予測分布の偏りに関する一考察 (パターン認識),” 電子情報通信学会論文誌. D-II, 情報・システム, II-パターン処理, vol.87, no.2, pp.407–415, feb 2004.
- [62] 風間淳一, ステインデ・サーガ, 黒田 航, 村田真樹, 鳥澤健太郎, “大規模分布類似度計算のためのベイズ手法を用いた新しい類似尺度,” 情報処理学会論文誌, vol.52, no.12, pp.3349–3362, dec 2011.
- [63] 木谷友哉, 澤 悠太, 柴田直樹, 安本慶一, 伊藤 実, “運転者に対する交通安全支援のための指向性アンテナおよび車車間通信を用いた歩行者の位置推定手法,” 情報処理学会論文誌数理モデル化と応用 (TOM), vol.7, no.2, pp.74–85, nov 2014.
- [64] 豊田敏裕, 本庄秀至, 中内茂樹, “シーン統計量に対するベイズ推定としての色恒常性 (バイオサイバネティックス, ニューロコンピューティング),” 電子情報通信学会論文誌. D, 情報・システム, vol.89, no.9, pp.2101–2112, sep 2006.
- [65] 深澤圭太, 石濱史子, 小熊宏之, 武田知己, 田中信行, 竹中明夫, “条件付自己回帰モデルによる空間自己相関を考慮した生物の分布データ解析,” 日本生態学会誌, vol.59, no.2, pp.171–186, jul 2009.
- [66] 福山恵子, 樋口知之, 魚住禎司, 河野英昭, 湯元清文, “初動が緩慢な波動現象開始時点の精密同定: Pi2 型地磁気脈動オンセットタイムの決定法 (デジタル信号処理),” 電子情報通信学会論文誌. A, 基礎・境界, vol.87, no.12, pp.1502–1510, dec 2004.

- [67] 小林潔司, 藤高勝己, “合理的期待形成を考慮した経路選択モデルに関する研究,” 土木学会論文集, vol.1993, no.458, pp.17–26, 1993.
- [68] F. Dieter, H. Jeffrey, L. Lin, S. Dirk, and B. Gaetano, “Bayesian filtering for location estimation,” IEEE pervasive computing, vol.2, no.3, pp.24–33, 2003.
- [69] C. Corinna and V. Vladimir, “Support-vector networks,” Machine Learning, vol.20, no.3, pp.273–297, 1995.
- [70] 伊波 靖, 高良富夫, “危険なシステムコールに着目した windows 向け異常検知手法,” 情報処理学会論文誌, vol.50, no.9, pp.2173–2181, sep 2009.
- [71] 山崎俊彦, 大島辰之輔, 相澤清晴, “各クラスへの中間出力値を用いた多クラス認識のための確信度処理,” 映像情報メディア学会誌, vol.69, no.8, pp.J257–J260, 2015.
- [72] T. Maeda, T. Yamasaki, and K. Aizawa, “Multi-stage object classification featuring confidence analysis of classifier and inclined local naive bayes nearest neighbor,” 2014 IEEE International Conference on Image Processing (ICIP), pp.5177–5181, Oct. 2014.
- [73] 小林正幸, 小西康夫, 石垣博行, “強化学習を用いたサポートベクターマシンのパラメータ最適化,” 日本機械学会論文集 C 編, vol.71, no.710, pp.2989–2995, 2005.
- [74] 青木一真, 黒柳 奨, クグレマウリシオ, ヌグロホ アントサトリヨ, 岩田 彰, “Confident margin を用いた svm のための特徴選択手法,” 電子情報通信学会論文誌. D-II, 情報・システム, II-パターン処理, vol.88, no.12, pp.2291–2300, dec 2005.
- [75] V.V. N., Statistical learning theory, 1 edition, Wiley, Sept. 1998.
- [76] G. Isabelle, W. Jason, B. Stephen, and V. Vladimir, “Gene selection for cancer classification using support vector machines,” Machine Learning, vol.46, no.1, pp.389–422, 2002.
- [77] A.W. Whitney, “A direct method of nonparametric measurement selection,” IEEE Transactions on Computers, vol.C-20, no.9, pp.1100–1103, Sept. 1971.
- [78] T. Marill and D. Green, “On the effectiveness of receptors in recognition systems,” IEEE Transactions on Information Theory, vol.9, no.1, pp.11–17, Jan. 1963.

- [79] M. Farid and B. Lorenzo, "Classification of hyperspectral remote sensing images with support vector machines," *IEEE Transactions on geoscience and remote sensing*, vol.42, no.8, pp.1778–1790, 2004.
- [80] 鈴木嵩史, ワン ユ, 加藤ジェーン, 間瀬健二, "クロスデータセットにおける行動識別器の学習アルゴリズム," *電気学会論文誌C (電子・情報・システム部門誌)*, vol.135, no.12, pp.1574–1582, 2015.
- [81] Y. Zhou, M. Kantarcioglu, and B. Thuraisingham, "Self-training with selection-by-rejection," *2012 IEEE 12th International Conference on Data Mining*, pp.795–803, Dec. 2012.
- [82] C. Chih-Chung and L. Chih-Jen, "LIBSVM: A library for support vector machines," *ACM Transactions on Intelligent Systems and Technology*, vol.2, pp.27:1–27:27, 2011.
- [83] 吉谷仁志, 黄瀬浩一, 松本啓之亮, "サポートベクトルマシンを用いた新聞記事からのプロフィール情報抽出," *電気学会論文誌C (電子・情報・システム部門誌)*, vol.124, no.11, pp.2260–2266, 2004.
- [84] 山崎高弘, 常盤欣一郎, "テキストの特徴とhtml構造を利用したweb文書の読みやすさ評価方法," *電気学会論文誌C (電子・情報・システム部門誌)*, vol.132, no.9, pp.1524–1532, 2012.
- [85] 村上龍希, 景山陽一, 西田 眞, 白澤洋一, "人物写真の選定支援を目的とした顔画像における白飛びの準定量評価," *電気学会論文誌C (電子・情報・システム部門誌)*, vol.133, no.11, pp.2098–2109, 2013.
- [86] 梶原祐輔, 中村宗広, 木村春彦, 大藪多可志, "疾病予防のための生体情報と気象情報を用いた気分子測システム," *電気学会論文誌C (電子・情報・システム部門誌)*, vol.135, no.6, pp.570–579, 2015.
- [87] 吉川雅博, 三河正彦, 田中和世, "筋電位を利用したサポートベクターマシンによる手のリアルタイム動作識別," *信学論 (D)*, vol.92, no.1, pp.93–103, jan 2009.
- [88] C.-W. Hsu and C.-J. Lin, "A comparison of methods for multiclass support vector machines," *IEEE Transactions on Neural Networks*, vol.13, no.2, pp.415–425, March 2002.
- [89] 新美礼彦, 小西 修, "ストリームカーネルマシンによるパラレルブースティング," *情報処理学会論文誌. データベース*, vol.2, no.4, pp.13–23, dec 2009.

- [90] 三浦 純, 森田英夫, ヒルドミヒヤエル, 白井良明, “Svm による物体と位置の視覚学習に基づく屋外移動ロボットの位置推定,” 日本ロボット学会誌, vol.25, no.5, pp.792–798, 2007.
- [91] 桂 浩章, 三浦 純, ヒルドミヒヤエル, 白井良明, “季節や天候の変化にロバストな視覚認識に基づく移動ロボットの屋外ナビゲーション,” 日本ロボット学会誌, vol.23, no.1, pp.75–83, 2005.
- [92] 千田陽介, 木室義彦, 財部修平, 長谷川勉, “配置不明なタグを用いた自己位置推定,” 電気学会論文誌C (電子・情報・システム部門誌), vol.128, no.7, pp.1212–1221, 2008.
- [93] 小中英嗣, “機械学習に基づく離散値入力制御系に対する制御器設計手法,” 電気学会論文誌C (電子・情報・システム部門誌), vol.132, no.6, pp.897–906, 2012.
- [94] M. Stefan and G.D. M, “An experimental study on pedestrian classification,” IEEE transactions on pattern analysis and machine intelligence, vol.28, no.11, pp.1863–1868, 2006.
- [95] 中山 優, 小林秀幸, 中井一文, 江崎修央, 山端直人, 糀谷 齊, 杉浦彰彦, “山間部における rssi を用いた進行ルート推定システムの向上,” システム制御情報学会論文誌, vol.26, no.2, pp.84–85, 2013.
- [96] 西岡壮大, 中井一文, 江崎修央, 杉浦彰彦, “受信信号強度を利用した猿の位置推定,” 電子情報通信学会技術研究報告. PRMU, パターン認識・メディア理解, vol.114, no.356, pp.49–54, dec 2014.
- [97] 北原司, 中井一文, 山本陽祐, 木村佳嗣, 江崎修央, 山端直人, “深度情報を利用した害獣の自動計数システムの開発,” 情報科学技術フォーラム講演論文集, vol.14, no.4, pp.577–580, aug 2015.
- [98] 伊藤将章, 中井一文, 江崎修央, 山端直人, 杉浦彰彦, “検知システムの長期測定結果に基づいた襲来規模の分析,” 電子情報通信学会総合大会講演論文集, vol.2015, no.2, p.590, feb 2015.
- [99] W. Kamin, K. Chris, and C. David, “A practical evaluation of radio signal strength for ranging-based localization,” ACM SIGMOBILE Mobile Computing and Communications Review, vol.11, no.1, pp.41–52, Jan. 2007.

- [100] 趙 漢居, 川崎宏治, 垣内洋平, 岡田 慧, 稲葉雅幸, “電波強弱監視に基づく複数無線モジュールの順次投下と環境地図作成を行う小型飛行ロボットの実現,” 日本ロボット学会誌, vol.32, no.7, pp.643–650, 2014.
- [101] 日坂翔馬, 三浦俊祐, 上條俊介, “実交差点における受信電波強度 (rssi) を用いた移動物体検知,” 生産研究, vol.66, no.2, pp.161–167, 2014.
- [102] 立石和也, 井家上哲史, “Rssi 方式における減衰定数の位置推定時決定手法,” 電気学会論文誌 C (電子・情報・システム部門誌), vol.128, no.10, pp.1509–1515, 2008.
- [103] 川村俊蔵, 伊谷純一郎 (編), サル 社会学的研究, 第2版, 中央公論社, 1967.

著者発表論文

論文誌への発表

- [1] Nakai K, Ezaki N, Sugiura A (2017) Prediction of the Appearance of Monkeys Based on Environmental Conditions. IEEJ Transactions on Electrical and Electronic Engineering, Vol.12 No.1, 132-139

国際会議への発表

- [1] Nakai K, Sugiura A, Ezaki N (2014) Prediction of the Appearance of Monkeys Based on Weather Data and Wireless Sensing Network. The seventh International Conference on Mobile Computing and Ubiquitous Networking (ICMU2014) 177-181

その他の発表

- [1] 柴田頼紀, 杉浦彰彦, 小林秀幸, 中井一文, 江崎修央, 山端直人, 糀谷齊 (2012) 環境認識型ネットワークを利用した信頼性の高い害獣検知システム. 情報処理学会全国大会講演論文集 2012(1),259-260
- [2] 中山優, 小林秀幸, 中井一文, 江崎修央, 山端直人, 糀谷齊, 杉浦彰彦 (2012) 山間部における RSSI を用いた進行ルート推定システムの提案. 情報処理学会全国大会講演論文集 2012(1),415-417
- [3] 石倉功規, 杉浦彰彦, 小林秀幸, 中井一文, 江崎修央, 山端直人, 糀谷齊 (2012) 接近検知無線システムを利用した猿の出現予測手法の提案. 情報処理学会全国大会講演論文集 2012(1),413-414
- [4] 中山優, 小林秀幸, 中井一文, 江崎修央, 山端直人, 糀谷齊, 杉浦彰彦 (2013) 山間部における RSSI を用いた進行ルート推定システムの向上. システム制御情報学会論文誌 26(2),84-85
- [5] 伊藤将章, 杉浦彰彦, 中井一文, 江崎修央 (2013) 猿出現予測システムの性能向上. 電子情報通信学会総合大会講演論文集 D-20-8

- [6] 岡崎正継, 中井一文, 江崎修央, 杉浦彰彦 (2013) 集落における猿の検知および住民への表示システム. 電子情報通信学会総合大会講演論文集 B-19-25
- [7] 中井一文, 西岡壮大, 岡崎正継, 江崎修央, 杉浦彰彦 (2013) 1 RSSI の確率地図を用いた猿の位置推定手法の検討. 電子情報通信学会総合大会講演論文集 B-19-4
- [8] 西岡壮大, 中井一文, 江崎修央, 杉浦彰彦 (2014) 受信信号強度を利用した猿の位置推定. 電子情報通信学会技術研究報告 PRMU, パターン認識・メディア理解 114(356),49-54.
- [9] 中井一文, 山本陽祐, 木村佳嗣, 江崎修央, 山端直人 (2015) 深度情報を利用した害獣の自動計数システムの検討. 情報処理学会研究報告 2015(CVIM-196),1-5.
- [10] 山本陽祐, 中井一文, 江崎修央, 山端直人 (2015) 深度情報を利用した害獣の自動カウント. 電子情報通信学会総合大会講演論文集 D-12-61
- [11] 北原司, 中井一文, 山本陽祐, 木村佳嗣, 江崎修央, 山端直人 (2015) 深度情報を利用した害獣の自動計数システムの開発. 情報科学技術フォーラム講演論文集 14(4),577-580
- [12] 伊藤将章, 中井一文, 江崎修央, 山端直人, 杉浦彰彦 (2015) 猿検知システムの長期測定結果に基づいた襲来規模の分析. 電子情報通信学会総合大会講演論文集 B-18-43

付録A 使用機器の諸元

表 A.1: 猿用発信機の諸元

Table A.1 Specification of the transmitter for monkeys

Item	Specification
Standard	ARIB STD-T99
Tx Frequency	CH-1: 142.94MHz CH-2: 142.95MHz CH-3: 142.96MHz CH-4: 142.97MHz CH-5: 142.98MHz
Modulation	MSK 1,200bps
Operating Temperature	-20°C to +60°C
Tx Time	Less than 1s in 5s
Tx Power	Less than 10mW
Battery	TL-5920
Consumption Current	Tx: Less than 16mA Sleep: Less than 10μA
Antenna	Wire Antenna
Belt	600mm × 25mm
External Dimensions	72mm × 40mm × 32 mm
Weight	135g

表 A.2: 猿用受信機の諸元

Table A.2 Specification of the receiver for monkeys

Item	Specification
Rx Frequency	CH-1: 142.94MHz CH-2: 142.95MHz CH-3: 142.96MHz CH-4: 142.97MHz CH-5: 142.98MHz
Radio Type	F2D: MSK 1,200bps
Sensitivity	-120dBm
Antenna Input	Impedance 50Ω BNC
Input Voltage	DC5.5 to 13.5V
Consumption Current	Less than 70mA
Operation	RESET switch
Indicator	POWER LED (red)
External Dimensions	100 × 100 × 40 mm
Weight	200g
Operating Temperature	-10°C to +55°C
Store Temperature	-20°C to +65°C