

Современные технологии идентификации лиц

Дмитрий Швецов

В статье рассматривается один из методов идентификации лиц с использованием радиолокации. В данном методе используется радиолокация непрерывного действия с частотной модуляцией, несущей частотой 61 ГГц и полосой пропускания 6 ГГц. Данные, полученные в результате радиолокационного сканирования лиц людей, передают в свёрточную нейронную сеть (CNN). В этой сети сигналы, полученные от нескольких антенных элементов, параллельно синтезируются, чтобы преобразовать радиолокационные сигналы в изображение, которое является входной формой для CNN.

Вступление

На протяжении многих десятилетий радарные датчики использовались для оценки скорости и расстояния до различных целей. В последние годы радарные датчики стали применять гораздо шире в смежных отраслях промышленности. Например, радарные датчики стали устанавливать на транспортные средства и использовать для автономного вождения, а также использовать их для сбора биометрической информации, такой как частота дыхания человека и частота сердечных сокращений. Кроме того, радарные датчики также используются для наблюдения за людьми в помещениях. В отличие от камер распознавания лиц, радарный датчик не меняет своих характеристик даже в полной темноте. Кроме того, есть ряд преимуществ использования таких датчиков с точки зрения конфиденциальности и защиты информации по сравнению с системами распознавания лиц на основе изображения.

Основное требование к радарным датчикам — это быстродействие и миниатюризация самого радара. Поскольку в рассматриваемой системе используется датчик с высокочастотным диапазоном в области миллиметровых волн, вопрос размеров был решён. Например, сейчас Google устанавливает в свой смартфон радиолокационный чип Soli и использует его для распознавания жестов. Кроме того, были проведены исследования по определению положения людей внутри транспортного средства с помощью малогабаритного сверхширокополосного (СШП) радиолокационного датчика.

В этой статье рассмотрим исследование метода идентификации человеческих лиц путём их сканирования радарными датчиками и метода машинного обучения на основе собранных данных. Метод распознавания лиц на основе радиочастотного сканирования может компенсировать недостатки метода распознавания лиц на основе изображений. В этом исследовании используется малогабаритный радиолокационный датчик с частотной модуляцией непрерывного излучения (FMCW) с несущей частотой 61 ГГц и полосой пропускания 6 ГГц. Поскольку метод FMCW предполагает высокую степень сжатия импульсов, он имеет лучшие показатели по дальности и точности измерений, чем другие методы с использованием частотной модуляции. Сначала проводится сбор данных с радиолокационных датчиков при сканировании лиц нескольких человек. Затем для идентификации лиц проводится обучение свёрточной нейронной сети (CNN) на основе собранных и обработанных данных. Поскольку CNN использует изображение в качестве входного форма-

та, необходимо преобразовать радиолокационные сигналы с датчиков в форму изображения. Для этого сигналы, полученные от нескольких радиочастотных датчиков, развёртывают параллельно, чтобы сформировать цельное изображение для дальнейшего обучения нейросети. Затем проводится анализ эффективности классификации сети, обученной с помощью данных «радиолокационных изображений». Помимо этого, проводится анализ радиолокационных данных для тех случаев, когда люди носят медицинские маски. Это важно для обучения сети в корректной классификации в зависимости от того, надета маска или нет.

Конфигурация стенда для радиолокации в диапазоне 61 ГГц

При исследовании радиолокационного метода для распознавания лиц людей был использован радиолокационный датчик миллиметрового диапазона BTS 60. На рис. 1 показана блок-схема радиолокационного датчика FMCW на 61 ГГц. На рисунке: передающая антен-

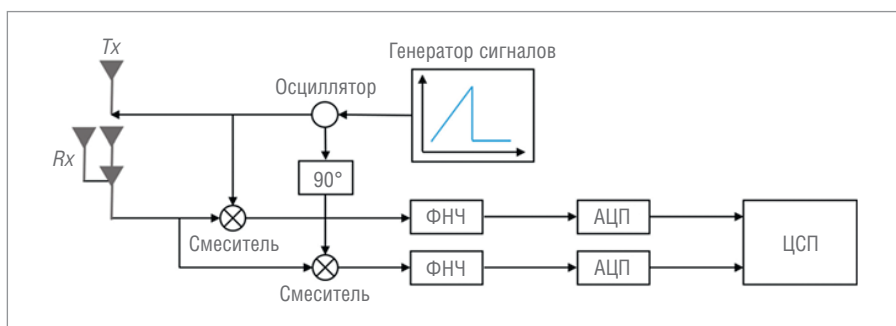


Рис. 1. Структурная схема радиолокационной системы FMCW 61 ГГц

на (Tx), осциллятор, фильтр низких частот (ФНЧ), АЦП и цифровой сигнальный процессор (ЦСП) относятся к передающей антенне, а к приёмной антенне: приёмные антенны (Rx), осциллятор, ФНЧ, АЦП и ЦСП, подключённые через смеситель. Как показано на рисунке, радар имеет одну передающую антенну и три элемента приёмной антенны (т.е. 1×3 антенную систему). Два элемента приёмной антенны размещаются горизонтально и два вертикально, при этом один антенный элемент используется совместно. Расстояние между соседними антенными элементами равно $c/2fc$, где c – скорость света, а fc – несущая частота 61 ГГц.

Этот радар передаёт форму волны, частота которой линейно возрастает, как изображено в генераторе сигнала на рис. 1. Этот датчик FMCW увеличивает полосу пропускания B на 6 ГГц на 12,5 мс. Таким образом, разрешение по дальности Δr становится равным 2,5 см (т.е. $\Delta r = c/2B$). Кроме того, в этом исследовании за период формы волны взят один кадр.

ЭКСПЕРИМЕНТАЛЬНАЯ УСТАНОВКА ДЛЯ СБОРА ДАННЫХ С РАДИОЛОКАЦИОННЫХ ДАТЧИКОВ

С помощью радара, описанного в предыдущем разделе, были проведены эксперименты по сбору данных. Как показано на рис. 2, эксперименты проводились с радаром, расположенным на расстоянии в $R = 30$ см от центра лица. Данные радиолокационных датчиков были собраны при соблюдении одинаковых экспериментальных условий для трёх разных людей. Кроме того, были получены радиолокационные данные лиц каждого человека в медицинской маске и без. В общей сложности было собрано 12 000 кадров радиолокационных сигналов для трёх испытуемых с медицинскими масками и без них.

ГЕНЕРАЦИЯ ГРАФИЧЕСКИХ ДАННЫХ ДЛЯ ОБУЧЕНИЯ CNN

В целом CNN – это широко используемая нейросеть для классификации нескольких изображений. Следовательно, важно преобразовать радиолокационные сигналы в формат изображения, соответствующий требованиям для входных данных CNN. В этом случае применяется метод синтеза радиолокационных сигналов из нескольких параллельных каналов. Другими словами, данные входного изображения:

$$X = \begin{bmatrix} |F(L_1[n])| \\ |F(L_2[n])| \\ \dots \\ |F(L_l[n])| \end{bmatrix}, \quad (1)$$

где $F(L_i[n])$ обозначает результат быстрого преобразования Фурье (БПФ) выходного сигнала ФНЧ с временной дискретизацией i_{th} ($i = 1, 2, \dots, l$) в приёмном канале.

На рис. 3 обобщённо показан процесс создания изображения из сигналов радара. В результате испытаний при ис-

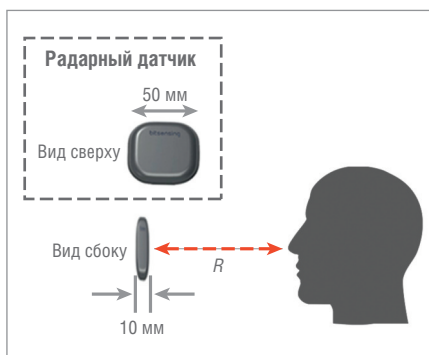


Рис. 2. Экспериментальная установка для сбора данных с радиолокационных датчиков

пользовании трёхэлементной приёмной антенны и 2^{10} точек БПФ было сгенерировано 3 изображения по 1024 пикселя с соответствующей размерностью. В случае применения метода радиолокации непрерывного излучения с частотной модуляцией (FMCW) эффективнее использовать сигналы, адаптированные к БПФ, а не сигналы, «привязанные» к временной шкале, поскольку расстояние между радаром и целью может быть легко рассчитано по изменению частоты сигнала.

Кроме того, для обучения нейросети используется не полное изображение X , а только изображение лица. Таким образом, из общего изображения «вырезается» область лица, которая выражается в X_R на рис. 3. В нашем примере на испытательном стенде размер X_R означает, что в диапазоне от $(R - 12,5$ см) до $(R + 12,5$ см) определяется, как нужная нам, область изображения лица.

Другими словами, только 12,5 см слева и 12,5 см справа изображения от текущего положения лица (R) используется в качестве входных данных нейросети.

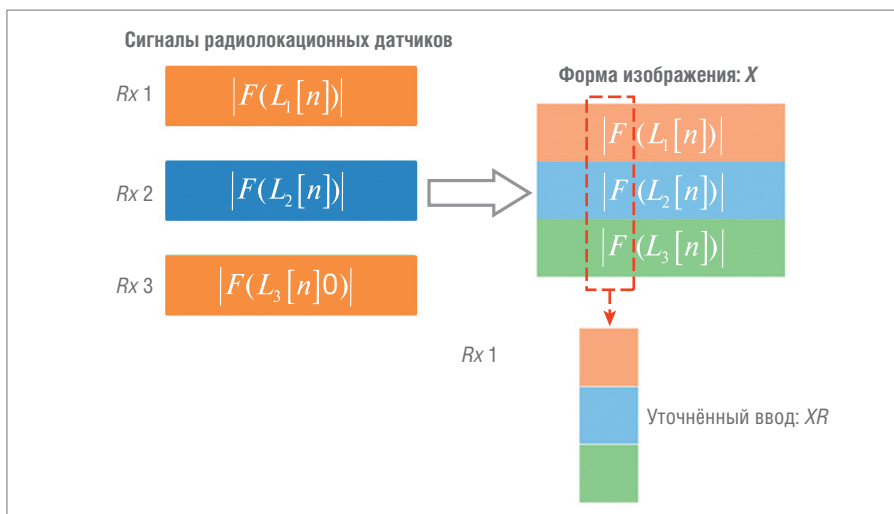


Рис. 3. Генерация входных данных для обучения CNN

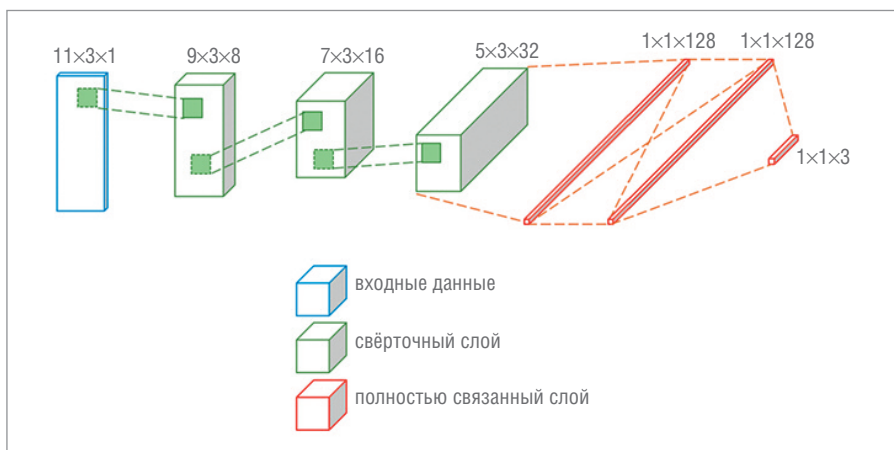


Рис. 4. Структура CNN, используемая для идентификации лица

Идентификация лица с помощью CNN

Структура используемой CNN модели приведена на рис. 4. В данном случае это три свёрточных слоя и три полностью связанных слоя в сети. В каждом свёрточном слое 3×3 выполняются свёртки, за которыми следует активация выпрямленного линейного блока и нормализация данных. Затем элементы последнего свёрточного слоя сглаживаются двумя полностью связанными слоями. После этих двух полностью соединённых слоёв добавляются отфильтрованные слои с коэффициентом отсева 0,25 для предотвращения переобучения. Последний полностью связанный слой даёт вектор размера $1 \times 1 \times 3$ на основе функции softmax. Наконец, мы рассчитываем потери модели с помощью кросс-энтропии. Подробные гиперпараметры модели CNN следующие: алгоритм импульса стохастического градиентного спуска используется для обновления параметров, размер мини-пакета установлен на 8, скорость обучения установлена на 10^{-4} , а эпоха установлена на 20.

Эта сеть обучается с помощью радиолокационных сигналов, полученных, когда трое испытуемых не были в масках. Чтобы предотвратить конкретное смещение классификатора, сеть обучается с использованием только 70% всех данных, а не всех данных, которые мы получили. Кроме того, 15% всех данных используется для набора проверки. Затем эффективность классификатора проверяется с использованием оставшихся 15% данных, и результаты классификации приводятся в табл. 1.

В нашем наборе данных радара средний результат идентификации для трёх лиц составляет около 98,7%. По сравнению с классификатором на основе глубокой нейронной сети (DNN), эффективность классификации в среднем выше.

Таблица 1
Результаты идентификации лиц для трёх испытуемых

Фактический класс / оценочный класс	Лицо 1, %	Лицо 2, %	Лицо 3, %
Лицо 1	97,3	2,7	0
Лицо 2	1,3	98,7	0
Лицо 3	0	0	100

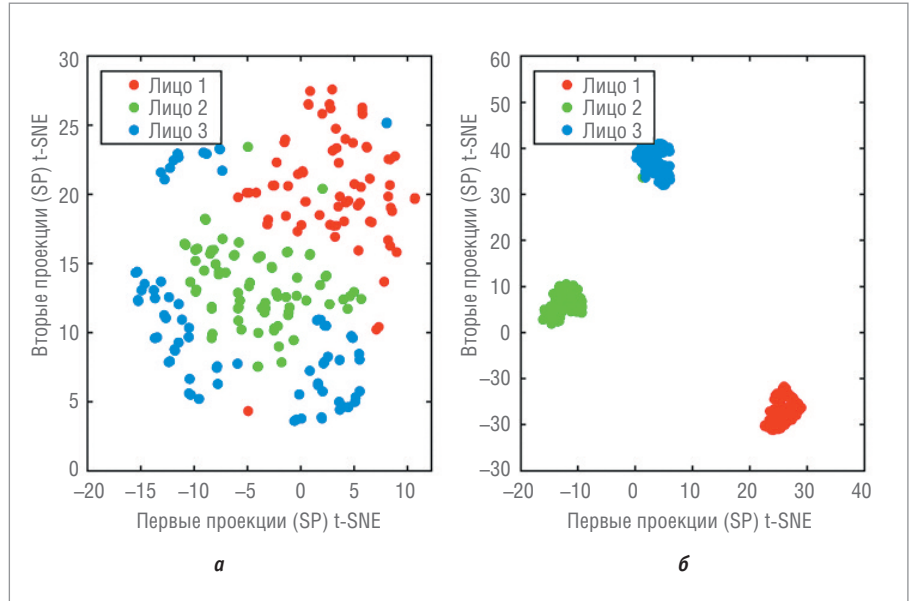


Рис. 5. Результаты применения алгоритма t-SNE

Кроме того, мы переобучаем CNN, добавляя данные радара, полученные, когда каждый субъект носил хлопковую маску, а характеристики классификации недавно обученной сети приведены в табл. 2.

При сравнении результатов идентификации табл. 1 и 2 средняя точность классификации снизилась на 0,9 процентных пункта. Эффективность метода распознавания лиц на основе радара не сильно ухудшилась в зависимости от того, была ли надета маска или нет.

Мы также проверяем эффективность классификации модели CNN с помощью метода визуализации данных высокой размерности, представляющего каждую точку данных в двух- или трёхмерном пространстве. Этот способ является модификацией метода стохастического вложения соседей (t-SNE). Алгоритм t-SNE может визуализировать данные большой размерности, также хорошо работает при уменьшении размерности. Алгоритм сохраняет подобие исходных данных, даже если размерность уменьшена до двух или трёх.

Таблица 2
Результаты идентификации лиц для трёх субъектов (включая данные радара, полученные, когда субъекты были в масках)

Фактический класс / оценочный класс	Лицо 1, %	Лицо 2, %	Лицо 3, %
Лицо 1	96,7	2	1,3
Лицо 2	0,7	96,7	2,7
Лицо 3	0	0	100

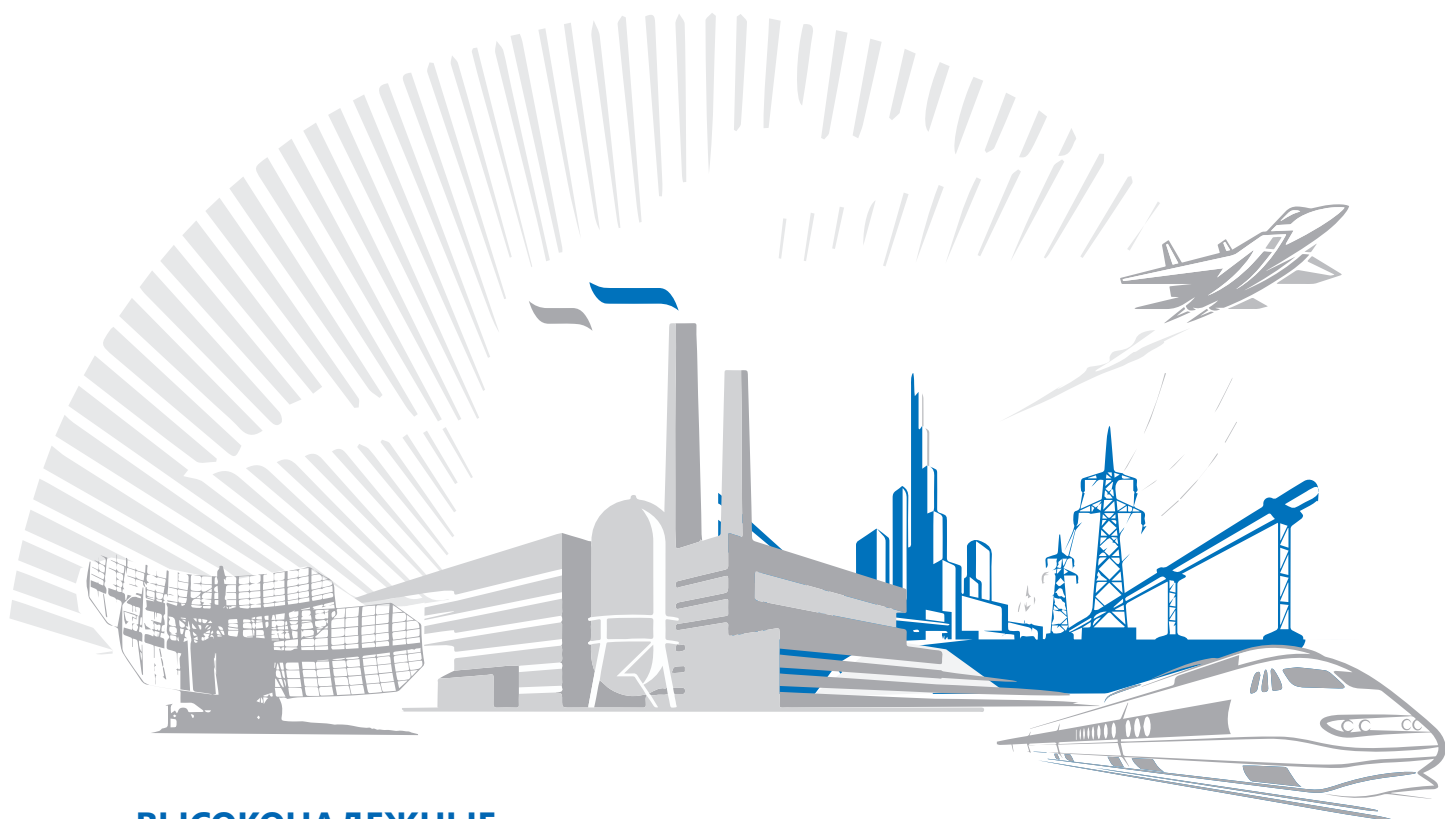
Следовательно, с помощью алгоритма t-SNE можно наблюдать сходство исходных данных на двумерной диаграмме рассеяния.

Результаты работы алгоритма t-SNE показаны на рис. 5. Там приведён результат применения алгоритма t-SNE к входному изображению XR. Как видно из рисунка, образцы из каждого класса не сгруппированы вместе, а широко распределены. Результат применения алгоритма t-SNE к выходному вектору, который получается из последнего слоя предложенной модели CNN, также представлен на рис. 5б. По сравнению с рис. 5а, внедрённые данные каждого класса расположены близко друг к другу. Таким образом, предлагаемая модель классификации эффективно извлекает признаки из входных данных лица.

ЗАКЛЮЧЕНИЕ

В статье приведён пример, как с помощью свёрточной модели CNN различать человеческие лица на основе данных радиолокационного сигнала миллиметрового диапазона. Вполне ожидаемо, что этот метод распознавания лиц на основе радиоволн может компенсировать недостатки метода распознавания лиц на основе изображений. Точность распознавания лиц с помощью CNN составляет > 98%. Кроме того, точность распознавания лиц в масках не ухудшается при использовании методов радиолокации. ●

Автор – сотрудник
фирмы ПРОСОФТ
Телефон: (495) 234-0636
E-mail: info@prosoft.ru



ВЫСОКОНАДЕЖНЫЕ МОДУЛЬНЫЕ ПРЕОБРАЗОВАТЕЛИ ДЛЯ ЖЕСТКИХ УСЛОВИЙ ЭКСПЛУАТАЦИИ



Серия HWS, HWS-A

- AC/DC-источники питания мощностью от 15 до 1560 Вт
- Ограниченная пожизненная гарантия
- Диапазон рабочих температур $-40...+70^{\circ}\text{C}$, конформное покрытие платы (модификация HD)
- Широкий диапазон выходных напряжений: от 3,3 до 60 В
- Работа в режиме пиковой мощности 300% в течение 5 с (модификация P)



Серия PFE, PFH

- AC/DC-преобразователи на плату от 300 до 1008 Вт
- Вход: 85–265 В AC, регулируемый выход: 12, 24, 48, 51 В DC
- Защиты от перенапряжений, перегрузки, перегрева
- Диапазон рабочих температур подложки $-40...+100^{\circ}\text{C}$
- Цифровое управление, обратная связь, поддержка PMBus™



Серия CN-A

- DC/DC-преобразователи на плату от 30 до 200 Вт
- Повышенная устойчивость к вибрациям, рекомендован для железнодорожного транспорта
- Диапазон рабочих температур $-40...+100^{\circ}\text{C}$ на подложке без снижения мощности
- Вход: 60–160 В DC или 14,4–36 В DC, выход: 5–24 В DC
- 5 лет гарантии



Серия HQA/GQA

- DC/DC-преобразователи на плату мощностью 85 и 120 Вт
- Ударные перегрузки до 50g
- Диапазон рабочих температур $-55 (-40)...+115^{\circ}\text{C}$
- Широкий диапазон входных напряжений: 9–40 В и 18–40 В DC
- Выходные напряжения: 5, 12, 15, 24, 28, 48 В DC
- КПД до 91,5%

