

Распознавание эмоционального состояния человека по его речи



«Только умея слушать и отвечать
можно быть хорошим собеседником.»

Ф. Ларошфуко

Мы начинаем цикл статей, посвященных эмоциональному состоянию человека, и его влиянию на компьютерные диалоговые системы. Эмоции и речь тесно взаимосвязаны и играют огромную роль в общении. В связи с этим, автоматическая и объективная диагностика эмоционального состояния человека по его речи представляет большой практический интерес. Исследованием данного феномена занимаются различные научные и коммерческие организации.

Что же такое эмоции? Существует множество определений. Вот некоторые из них. Эмоции - сильные психические состояния, связанные обычно с возбуждением или высоким уровнем энергии и дающие начало чувствам и страстям [1];. Также, чувствами или эмоциями называют переживание человеком своего отношения к тому, что он познает и делает, к другим людям и к самому себе [2];. Эмоции обычно бывают положительными или отрицательными. Удивление, эйфория, гнев, страх различаются по степени положительности или отрицательности. Эмоции дают нам информацию о том, как говорящий оценивает ситуацию, и какие ответные реакции можно ожидать от него. Необходимо уметь интерпретировать эмоции, уменьшать их интенсивность, направлять или преобразовывать одну эмоцию в другую. Выяснено, что мысли и эмоции влияют на дыхание, выражение лица, положение тела, тон и темп голоса. Голос выражает любые сильные эмоции, он принимает, по закону ассоциации [3], музыкальный мелодичный характер, меняясь по громкости и по качеству, то есть - по резонансу, тембру, высоте звука. Темп речи и её ритмическое (акцентное) членение с помощью пауз и логического ударения также имеют выразительное значение, помогающие уловить эмоции человека.

В результате проведённых исследований и анализов выявлено, что существует такое понятие как эмоциональный слух, т.е. способность распознавания эмоций по речи человека. Испытуемые разного возраста, пола и профессии показали существенные различия в правильности понимания эмоций — от 10 до 95%, так выявлено, что музыканты обладают более развитым эмоциональным слухом. Для определения эмоций важен эмоциональный фон испытуемого, т.е. какую эмоцию он сам переживает в данный момент. Для результативного и успешного диалога важны мониторинг состояния говорящего и отображение понимания этого состояния. Учёные университета Эль-Пасо (США) выбрали для изучения такое понятие, как уровень уверенности высказывания, т.е. насколько уверенно говорящий произносит то или иное высказывание. Анализ проведённых исследований, основанных на диалоге студентов и преподавателя, а затем студентов и компьютерной записи речи преподавателя, показал, что определить уровень уверенности высказывания легко, если в системе представлено высказывание, содержащее только одну основную идею. Тем не менее, в естественном языке высказывания могут быть очень длинными и редко хорошо и чётко сформулированными. Идея исследования заключалась в том, чтобы создать модель прогнозирования уровня уверенности. Высказывания, на которых практиковалась модель — это высказывания различных уровней уверенности, они взяты из речи носителей английского языка[5].

Уровень уверенности говорящего определяется тоном и высотой голоса:

- явно высокий — энтузиазм, радость, заинтересован и проявляет интерес;
- высокий, в широком диапазоне силы, тональности и высоты — гнев и страх, неуверенность;
- чрезмерно высокий, пронзительный — беспокойство;

- мягкий и приглушенный, с понижением интонации к концу каждой фразы — печаль, усталость;
- форсирование звука — напряжение, обман.
- в состоянии эмоционального возбуждения обычно возрастает сила голоса, изменяются его высота и тембр, но иногда сильное возбуждение может, наоборот, проявляться в уменьшении силы голоса (человек «шипит от ярости»).

Ученые Саутгемптонского университета (Великобритания) разработали компьютерные методы, позволяющие прогнозировать ответную эмоциональную реакцию говорящего. В ходе эксперимента было выявлено, что просодическая информация помогает в автоматическом определении степени раздражённости человека. Наиболее полную информацию о внутреннем психоэмоциональном состоянии человека может дать анализ его связной речи: расстановка логических ударений, скорость произнесения слов, конструкция фразы, наличие таких отклонений от нормы, как неуверенный или неверный подбор слов, обрывание фраз на полуслове, изменение слов, появление слов-паразитов, исчезновение пауз и т.д. [4]

Можно выделить:

- быстрая речь — очевидная взволнованность, страстное желание убедить или уговорить кого-то;
- медленная речь — высокомерие, усталость, угнетенное состояние;
- прерывистая речь — неуверенность;
- лаконичность и решительность речи — явная уверенность;
- заикание — напряженность или обман;
- нерешительность в подборе слов — неуверенность в себе или намерение внезапно удивить чем-то;
- появление речевых недостатков (повторение или искажение слов, обрывание фраз на полуслове) — несомненное волнение, но иной раз и желание обмануть;
- опускание речевых пауз — напряжение;
- слишком удлиненные паузы — незаинтересованность или несогласие;
- появление в речи пауз, заполняемых словами-паразитами — нерешительность и затруднение в выражении мысли, поиск выхода из положения;
- возрастание числа тривиальных наборов слов, проговариваемых быстрее, чем обычно, — эмоциональное возбуждение, напряжение;
- умолкание или скупость в словах — обида.

Человеческий голос, а также такие звуковые инструменты, как флейта, барабаны влияют на состояние мозга человека, а точнее, на его нейрологическую активность. Голосом можно контролировать и направлять эмоции другого человека, так, например, некоторые шаманы лишь при помощи голоса могут вводить людей в транс. В мозгу зафиксированы различные типы колебаний от медленного альфа ритма до высокочастотного бета ритма.

Выделяют следующие ритмы:

- Альфа-волны, 13 Герц - при нормальном состоянии, стрессе, состоянии тревоги.
- Бета-волны, 8 -12 Герц - при лёгкой релаксации, позитивных настроениях.
- Тета-волны, 4 - 7 Герц - глубокая релаксация, состояние сосредоточенного внимания.
- Дельта-волны, 1-3 Герц - глубокий сон, "ясные сновидения.

- Гамма-волны - до конца не исследованы.

Для развития стимулирующих, успешных диалоговых систем необходимы исследования по выявлению эмоций человека по его речи. Понимание эмоций другого человека важно как для общения между людьми, так и при взаимодействии человека с системами искусственного интеллекта. Автоматическое распознавание речи и прогнозирование эмоций говорящего нашло бы активное применение, например, в телекоммуникационной сфере и индустрии развлечений, что помогло бы избегать конфликтных ситуаций.

У всех людей эмоции проявляются по-разному. При определении эмоций по голосу, его тембру и интонации, необходимо учитывать индивидуальные особенности человека. Помимо того, человек может одновременно испытывать сразу несколько эмоций. Распознавание эмоционального состояния человека по его речи не простая задача.

[1] - www.zipsites.ru

[2] - www.big-library.info

[3] - www.galactic.org.

[4] - www.isca-speech.org.

[5] - Jaime C. Acosta and Nigel G. Ward «Responding to User Emotional State by Adding Emotional Coloring to Utterances»

Все материалы, размещенные на данном сайте, разрешены к публикации и печати на других ресурсах и печатных издания только при наличии письменного разрешения компании ООО "Речевые Технологии"