

Е.В.Бичева (6 курс, каф. САиУ), Б.И.Морозов, проф.

КОМПЛЕКСНАЯ СИСТЕМА РАСПОЗНАВАНИЯ РЕЧИ

В настоящее время достаточно большое внимание уделяется проблеме распознавания образов, звуков, видеоизображения и др. Одним из основных считается вопрос распознавания речи, которая является комплексным процессом и затрагивает не только акустические аспекты речеобразования (т.е. звуков, фонем, фраз), но и кинематику мимики и артикуляции. Данные исследования направлены на усовершенствование пользовательского общения с техническими средствами (сюда можно отнести и разработки «естественного» интерфейса в общении человека с электронными устройствами). Главным образом, отличие данных исследований состоит в том, что их результаты могут быть использованы и в области реабилитации слабослышащих пациентов для постановки речи.

Проблемой распознавания речи является ее многокомпонентность, т.е. сам звук произносимых фраз, движение лицевых мышц и слежение за артикуляцией в целом. Принимая во внимание вышеупомянутый комплексный метод, а также используя системный подход, перспективную компьютерную систему визуального контроля произношения и артикуляции можно представить в виде структурной схемы (рис.1).

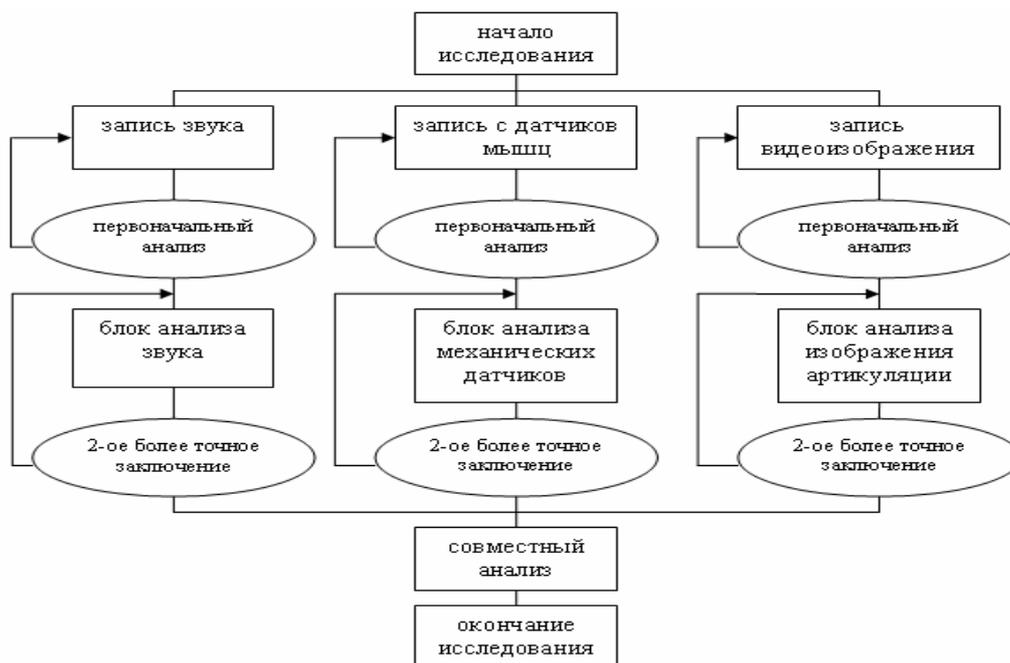


Рис 1. Структурная схема системы распознавания речи

Эта структурная схема состоит из трех параллельно действующих каналов. По первому каналу сначала ведется запись звука, далее производится первоначальный диагноз, затем анализ речи с использованием таблиц, графиков, диаграмм и др., после чего ставится второе, более точное заключение врача-логопеда. По второму каналу производится запись данных от механических датчиков, которые крепятся непосредственно на лицевые мышцы, далее по алгоритму аналогично с первым каналом. И, наконец, по третьему каналу осуществляется запись видеоизображения и производится анализ. В итоге производится совместный анализ.

Комплексная система распознавания речи, рассматриваемая в данной работе, имеет три основных компонента, которые настраиваются в соответствии с полученными результатами о качестве произношения. Единственный из этих трех компонентов, не подвергающийся никакому сравнительному анализу вследствие его наибольшей важности, – это анализатор звука произносимой речи. Однако возможны случаи использования данной комплексной системы в условиях ограничений (например, в условиях недостаточного финансирования и др.), когда необходимо сделать правильный выбор из двух оставшихся компонентов. Для этого удобно применять информационные оценки (методы управления внедрением нововведений, базирующиеся на использовании информационного подхода).

Итак, объект, рассматриваемый в данной статье, имел аналоги, которые, однако, не включали в свой состав анализаторы видеоизображения и сигналов с мимических датчиков. Таким образом, можно считать, что эти два последних элемента и есть нововведения. Т.е. можно провести анализ внедрения этих нововведений, это будет выражаться в следующей структурной схеме (рис. 2). Для оценки первого НВВ принято два критерия, для оценки второго – один критерий.

Первоначально относительная значимость первого НВВ возьмем выше, чем относительная значимость второго НВВ.

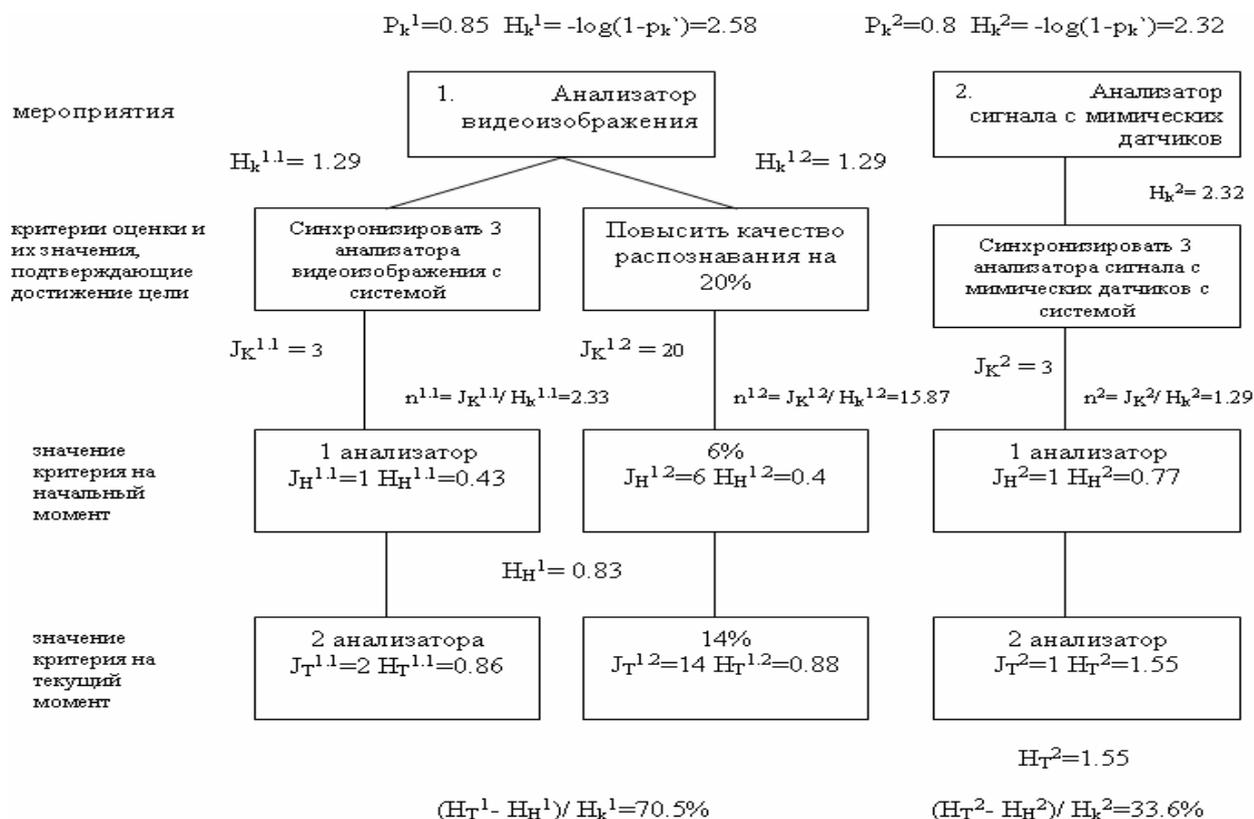


Рис. 2. Структурная схема анализа внедрения нововведений

Здесь p – вероятность недостижения цели, q – вероятность использования конкретного НВВ при реализации, достижении соответствующей подцели, H – значимость нововведения, $J=A_i/\Delta A_i$, A_i может интерпретироваться как число внедряемых единиц новой технологии, ΔA_i характеризует, с какой степенью точности нужно учитывать A_i в конкретных условиях, т.е. с помощью этого значения задаются единицы измерения, которые могут быть различными;

Таким образом, из приведенной выше схемы видно, что даже при учете изменения параметров J в процессе внедрения НВВ предпочтения не изменились, как была

первоначально относительная значимость первого НВВ выше, так и осталась выше; данная схема только подтверждает этот факт.

Итак, можно сказать, что применение системного подхода к такому комплексному исследованию позволит повысить качество распознавания речи в полтора-два раза.

Предполагается использовать полученные результаты в экспериментальных исследованиях и проверке эффективности комплексного метода для реабилитации слабослышащих пациентов, а также для управления автоматическими роботами.