

бесконтактные интерфейсы — компьютер управляемый взглядом

крик касперски, aka мышъх, a.k.a. nezumi, a.k.a. souriz, a.k.a. elraton, no-email

в последние годы на рынке появилось множество относительно дешевых бесконтактных устройств ввода, подчиняющихся... простому человеческому взгляду! ориентированные главным образом на людей с ограниченными возможностями, они, тем не менее, нашли применение в играх, смартфонах, цифровых фотоаппаратах и прочей бытовой технике. хотите узнать как работают бесконтактные интерфейсы и как соорудить подобное устройство самостоятельно? тогда — эта статья для вас!

введение

Зародившись в недрах военных лабораторий, бесконтактные интерфейсы наконец-то вырвались на оперативный простор потребительского рынка, потеснив мышь и отчасти клавиатуру. Однако, подавляющее большинство пользователей даже не подозревает о существовании подобных устройств в открытой продаже, между тем их цена стремительно падает, вплотную приближаясь к отметке в 100\$ (что сопоставимо со стоимостью хорошей клавиатуры с позолоченными контактами).

А ведь какой-то десяток лет назад на этой теме еще висел ярлык секретности и хотя гражданские институты разных стран раскрутили ее независимо от военных, в открытой печати принципы работы бесконтактных интерфейсов описаны очень поверхностно — с ошибками, белыми пятнами и недоговоренностями, с которыми приходится сталкиваться всякому, дерзнувшему реализовать подобное устройство самостоятельно. Немного забегая вперед, отметим, что для этого потребуется всего лишь web-камера среднего класса и умение программировать на любом языке, но мы несколько отвлеклись от темы...



Рисунок 1 компьютер, подчиняющийся взгляду, это уже не миф, а реальность, доступная каждому желающему!

как это работает?

Точная дата появления первых устройств управления взглядом до сих пор остается предметом острых дискуссий (перец, горчица и прочие приправы — по вкусу). По косвенным данным, опирающихся на успехи научно-технического прогресса в смежных областях, нижнюю границу оценок можно отнести к 70ым годам прошлого века, хотя сами принципы, положенные в основу интерфейсов подобного типа, были известны (и широко использовались в военных целях) еще во времена Октябрьской Революции.

Плакат "Ты записался добровольцем" — помните? Человек в красном смотрит на нас пронзающим взглядом где бы мы ни находились. Почему?! Да потому, что направление взгляда определяется положением зрачка, а человека на плакате зрачки расположены в самом центре, в результате чего создание любого наблюдателя интерпретирует ситуацию так, что дзен-будденовец смотрит прямо на него! И от этого взгляда никуда не скрыться! Даже если отойти на несколько шагов в сторону, нарисованные зрачки по-прежнему останутся в центре, порождая иллюзию, что картина следит за тобой!!! Довольно жуткое впечатление, особенно для неизбалованных телевидением людей.



Рисунок 2 "живой" плакат, перемещающий взгляд вслед за проходящими мимо него зрителями

Если же нарисованные зрачки смешены относительно центра глаза — это воспринимается как взгляд в сторону, направленный на вполне конкретный (или подразумеваемый) объект, местоположение которого легко вычислить, мысленно продолжив взгляд или нарисовав линию, соединяющую центр глаза с центром зрачка. Она и даст направление взгляда в аксонометрической или изометрической проекции. Это в том случае, если голова повернута в сторону рассматриваемого предмета, в противном случае, придется вносить соответствующие корректизы.



Рисунок 3 система управления взглядом должна уметь с максимальной точностью определять центры глаза и зрачка

Смещение зрачка относительно центра глаза (см. рис. 3) элементарно регистрируется видеокамерой с любым (разумным) линейным разрешением и легко обрабатывается даже тормозными микропроцессорами, а вот определить ориентацию головы алгоритмическим путем не так-то просто, но создатели первых систем наведения, используемых в реактивных истребителях, нашли весьма элегантное решение проблемы, вмонтировав видеокамеру непосредственно в шлем, одетый на пилота, в результате чего при повороте головы относительно положение камеры остается неизменным, поскольку камера поворачивается вместе с самой головой, ну а дальше, как говорится, дело техники. Закрепляем еще одну камеру под самолетом и синхронизуем ее движение с движением шлема. Зная расстояние от глаза до камеры и определив относительное смещение зрачка, вычисляем угол направления взгляда через простейшие тригонометрические функции (которые в процессорах первых поколений, были представлены в виде предвычисленных таблиц, ибо сопроцессоры тогда были довольно дорогими штуками).

Снимаем изображение с камеры под самолетом и находим объект, находящийся на заданном угловом расстоянии относительно оптической оси, после чего увеличиваем его и обрамляем прямоугольником, захватывая цель и автоматически отслеживая ее перемещение, после чего пилоту остается нажать большую красную кнопку типа гашетки и ракета со свистом накроет цель по полной программе. Просто, быстро, удобно, надежно и, главное, что руки свободы, в смысле заняты управлением самолетом. До этого момента экипаж самолетов подобного класса состоял как минимум из двух человек — один крутил штурвал, другой — стрелял по врагам, что увеличивало габариты кабины и отрицательно влияло на ТТХ всего летательного аппарата, которому в боевых условиях приходилось не только летать, но и падать.

Ну, самолеты — это понятно. А вот компьютеру подобная техника зачем?! Действительно, большинство пользователей крайне скептически относятся к бесконтактным интерфейсам типа голосового управления и прочих извращенных инноваций подобного типа. Они эффективно смотрятся на выставках, но выявляют огромное количество недостатков при попытках практического применения: низкая точность, ненадежность, большое количество ложных срабатываний, etc. Все это дискредитирует саму идею бесконтактных интерфейсов и публика, наигравшись с сырьими образцами, в спешке выброшенными на рынок, быстро утрачивает к ним всякий интерес, возвращаясь к своей любым мышам и клавиатурам. Между тем, управление взглядом не только быстро, удобно, но еще и чрезвычайно надежно! В правильно спроектированной и должным образом реализованной системе, процент ложных срабатываний стремится к минимуму.

И вот теперь настало самое время обсудить какими именно техническими средствами все это достигается и где их можно приобрести.

компьютер, управляемый взглядом

Заставить компьютер, слушаться взгляда, намного сложнее чем самолет, поскольку, во-первых, требуется обеспечить намного более высокую точность позиционирования курсора, а, во-вторых, в системе голова — камера — монитор, все компоненты могут занимать различные положения. Человек откинулся в кресле, повернул монитор под другим углом, переставил камеру, и... вся телеметрия идет коту под хвост, требуя перекалибровки.

На самом деле, все эти трудности вполне преодолимы. Начнем с того, что монитор представляет собой плоскость, пространственное положение которой известно системе, причем эта плоскость находится очень близко к человеку, настолько близко, что вступает в игру такое явление как параллакс — человек имеет два глаза, разнесенных приблизительно на 10 см, уставившихся в монитор, расположенный в где-то в полуметре. Допустим, центр поля зрения расположен строго посередине монитора (см. рис. 4). Тогда, (как нетрудно рассчитать) левый зрачок будет смотреть под углом 6° относительно линии, перпендикулярной плоскости монитора, а правый — под углом -6° , что в совокупности нам дает 12 градусов. Весьма весомая величина, которую нетрудно измерять даже камерой низкого разрешения!

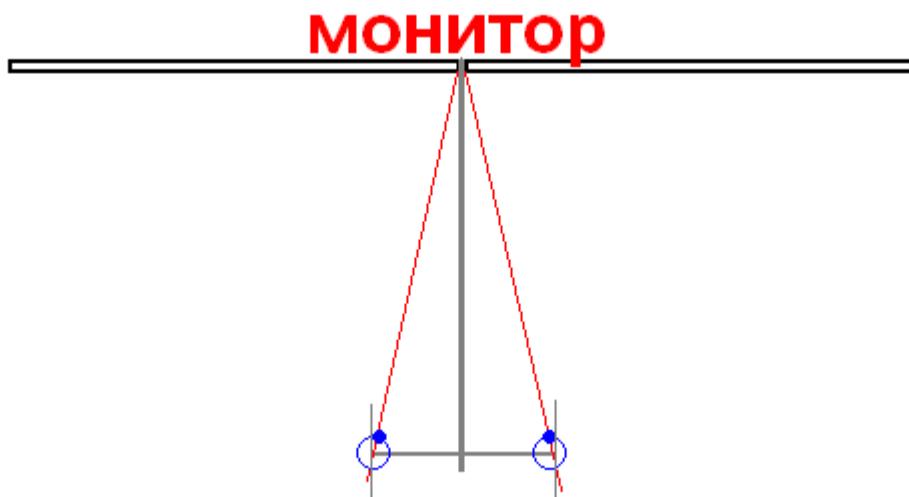


Рисунок 4 использование эффекта параллакса для определения точки, на которую направлен взгляд, фиксируемый жестко закрепленной на мониторе видеокамерой

Напротив, при взгляде на бесконечно удаленный объект (каким, например, является танк с высоты полета истребителя), разница углов зрения стремится к нулю и потому параллакс в бортовых системах наведения — плохой помощник, но зато при управлении компьютером он превращается в основное средство, нивелирующее положение головы, ведь направление взгляда в этом случае определяется через разницу смещений зрачков относительно центра взгляда, которое остается постоянным при любой ориентации головы, правда, чрезвычайно чувствительно к изменению расстояния между головой и монитором. Однако, расстояние до головы (в отличии от ее пространственной ориентации) определяется

элементарно — по угловому промежутку между глазами. При приближении к монитору, количество пикселей между ними будет возрастать, а при удалении — уменьшаться.

Измеряя линейное расстояние между глазами на снимке и смещение зрачков относительно центра, мы получаем самокалибрующуюся систему, не требующую облачения пользователя в шлем. Камера в этом случае закрепляется около монитора (сверху или сбоку — без разницы). Вращаясь вместе с монитором, она представляет единую систему, что позволяет выполнять все расчеты в абсолютной системе координат, привязанной к плоскости монитора (см. рис. 5). Естественно, мы должны учитывать геометрические размеры и текущее разрешения монитора, тесно взаимодействуя с операционной системой, но это уже детали.

Кстати говоря, LCD-мониторы в этом смысле намного более предпочтительны, поскольку лишены геометрических искажений и, что самое главное, размер изображения всегда совпадает с размером экрана. В CRT-мониторах последнее условие выполняется далеко не всегда и, подкручивая ручки настройки, пользователь может варьировать размер изображения в широких пределах, ужимая или растягивая его как по горизонтали, так и по вертикали, причем, ни видеокарта, ни операционная система об этом не имеют никакого представления и определяют эффективный размер экрана исключительно на основе заводских установок данного монитора, которые в общем случае не совпадают с действительными.



Рисунок 5 NaturalPoint trackIRT от компании Today Eye Control Technologies, Inc. — популярное средство, управления компьютера взглядом, основанное на параллаксе (сама фирма не разглашает технических деталей, но все секреты легко вскрываются дизассемблером)

типы систем управления взглядом

Самые совершенные (и, соответственно, самые дорогие) системы управления взглядом, представляют собой камеру, закрепленную на голове (чуть ниже уровня глаз) и снабженную датчиками пространственного положения, за счет которых обеспечивается свобода перемещения человека относительно монитора (он может не только сидеть, но и ходить), а также высокая точность позиционирования, достигающая нескольких пикселей, что позволяет рисовать картинки, работать с картографическими приложениями и программами видеонаблюдения, быстро "выхватывая" взглядом подозрительных людей из толпы, даже если перед нами находится целая дюжина громоздящихся мониторов.

Основным минусом подобных систем (помимо высокой стоимости) является низкая "помехоустойчивость". Даже если оператор полностью сосредоточен, глаз все равно совершает множество хаотичных движений, непроизвольно бегая по всему полю зрения, причем в момент нервного напряжения (или повышенной усталости) интенсивность хаотичных блужданий резко возрастает, а, поскольку, "скорострельность" видеокамеры не бесконечна, система "усредняет"

блуждания глаза и теряет управляемость, будучи не в состоянии отфильтровать сознательные перемещения от бессознательных.

Над решением этой проблемы ломают голову тысячи инженеров и ученых. Исследовательскому коллективу Фраунгоферского Института в Германии (Fraunhofer's Institute Industrial Engineering team) как будто бы удалось создать фильтр, основанный на сложных психофизических моделях и отсеивающих большое количество помех ("statistical average filter"), однако, практические перспективы его применения до сих пор неясны, а описание самих алгоритмов фильтрации недоступно, так что широкой публике остается только гадать, чем она, собственно говоря, и занимается. В распространенном пресс-релизе упоминается всего лишь тот факт, что цвет кнопки, на которую направлен взгляд, дважды изменяется для получения подтверждения от пользователя (подробности см. на новостном блоге: http://blog.digitalhomemag.com/page/digitalhome?entry=eye_controlled_interaction_for_your). Не самое лучшее решение, т. к. "мельканье" ошибочно распознанной кнопки тут же привлечет внимание пользователя, немедленно переводящего на нее взгляд. Даже небольшое количество ошибок, вызовет сильное "мельтешение" всего экрана, приводящее к быстрой утомляемости и, как следствие, снижению точности наведения.



Рисунок 6 здание Фраунгоферского Института в Германии в котором работает исследовательский коллектив по созданию бесконтактных интерфейсов, управляемых взглядом

Более дешевые системы используют камеру, закрепленную на мониторе, и работающую посредством параллакса. Точность наведения падает до десятков пикселей и потому для работы в Photoshop'e они уже непригодны, зато уверенно двигают мышку по кнопкам, меню и другим элементам управления. Пользователь может бродить по Сети, не отрываясь от чашки кофе, переворачивать страницы электронных книг, раскладывать пасьянс или даже играть в динамичные игры типа стрелялок, правда, с последними система справляется уже на пределе, особенно если игрок нервно дергает головой и всем остальным телом в придачу. Автокалиброка занимает доли секунды, но этого времени вполне достаточно для того, чтобы вынырнувший из-за угла монстр, отправил нас на свет иной.



Рисунок 7 авиационный симулятор, поддерживающий системы управления взгляда типа NaturalPoint trackIRT и другие, совместимые с ним

Другой недостаток — одновременная работа с несколькими мониторами, как правило, не поддерживается, свобода передвижений пользователя весьма ограничена (даже сильное отклонение в кресле дезориентирует систему), наконец, перед началом первого сеанса работы с системой каждый пользователь должен выполнять калибровку под себя, в процессе которой система зажигает в различных частях монитора мерцающие перекрестья, на которых необходимо сфокусировать взгляд, что занимает порядка десяти секунд. Вроде бы мелочь, но на персональных компьютерах коллективного использования она здорово напрягает. Впрочем, все эти недостатки с лихвой окупает низкая цена, находящаяся в районе сотни долларов. На западе весьма популярен программно-аппаратный комплекс "NaturalPoint trackIRT" от компании "Today Eye Control Technologies, Inc." со стоимостью в \$99 за базовый комплект, позволяющий перемещать мышиный курсор взглядом (<http://www.naturalpoint.com/trackir/> и http://www.naturalpoint.com/smartnav/corporate/press/03_20_01.html).

Изначально предназначенный для людей с ограниченными возможностями, он, тем не менее, пришелся по вкусу и нормальным пользователям, получив поддержку со стороны многочисленных производителей программного обеспечения (и, в первую очередь, разработчиков игр). Кстати говоря, уже выработан унифицированный интерфейс для подобных устройств и потому их можно использовать с любыми программами, которые только его поддерживают. До этого, каждой устройство могло работать только с ограниченным набором специально "заточенных" под эти цели приложений. Так что прогресс не стоит на месте и скоро системы управления взглядом из разряда новинок превратятся в привычные девайсы, мирно соседствующие с мышью и клавиатурой.

Однако, сто долларов — это достаточно большая сумма, которую не все готовы выложить на бесконтактный интерфейс, а потому самую низко бюджетную нишу занимают устройства, основанные на базе web-камеры низкого разрешения (в комплект, как правило, не входящей и приобретаемой отдельно), жестко закрепленной на мониторе и фиксирующей смещение зрачков, относительно некоторое точки, запоминаемой при калибровке. Точность позиционирования падает в разы, впрочем, для работы с виртуальной клавиатурой, этого вполне достаточно. Чтобы словить идею, достаточно посмотреть на "My Tobii" от компании "Eye Control System" (www.speechgeneratingdevices.com/tobiiitobii.html). Его можно встретить в различных реабилитационных центрах для парализованных людей или просто общественных

терминалах, владельцам которых надоело выковыривать остатки гамбургеров из-под клавиатуры. Бесконтактные интерфейсы в этом отношении намного более предпочтительны, и (что тоже немаловажно) абсолютно бесшумны.



Рисунок 8 виртуальная клавиатура бесконтактного интерфейса "My Тобii" от компании "Eye Control System", управляемая взглядом и ориентированная на людей с ограниченными возможностями

система управления взглядом — своими руками

При наличии Web-камеры (среднего или высокого класса) бесконтактный интерфейс можно реализовать и самостоятельно. Как уже говорилось, есть два пути — жестко закрепить камеру на голове (но тогда нам потребуется соорудить еще и датчики пространственной ориентации, что в домашних условиях сделать довольно затруднительно), либо расположить ее непосредственно на самом мониторе. Ну на мониторе, так на мониторе. Сказано — сделано!

Теперь мы можем запустить программу "грабежа" картинок и "фотографировать" глаза оператора с частотой порядка десятки кадров в секунду или немногим менее того (можно и быстрее, но не нужно). Что? Не получается?! И не получится. Темно слишком. Зрачки сильно расширены и практически полностью сливаются с радужной оболочкой. Даже если камера обладает нехилым разрешением, смещение зрачков относительно центра глаз определяется весьма неуверенно. Фактически, мы можем распознать лишь крайние ситуации — взгляд в центр, вверх/вниз, вправо/влево и в каждый из четырех углов. Хм, а что?! Для листания книги этого вполне достаточно! Система, отслеживая положение зрачков, автоматически прокручивает страницу, опуская тест вниз или возвращаясь наверх, повинуясь взгляду читателя.

Чтение электронных книг при этом значительно упрощается, особенно если руки в это время заняты чем-то другим (не будем уточнять конкретно чем, ведь помимо технической литературы еще существует эротика и масса других интересных направлений, вызывающих острую тоску по отсутствию хвоста или другого агрегата).

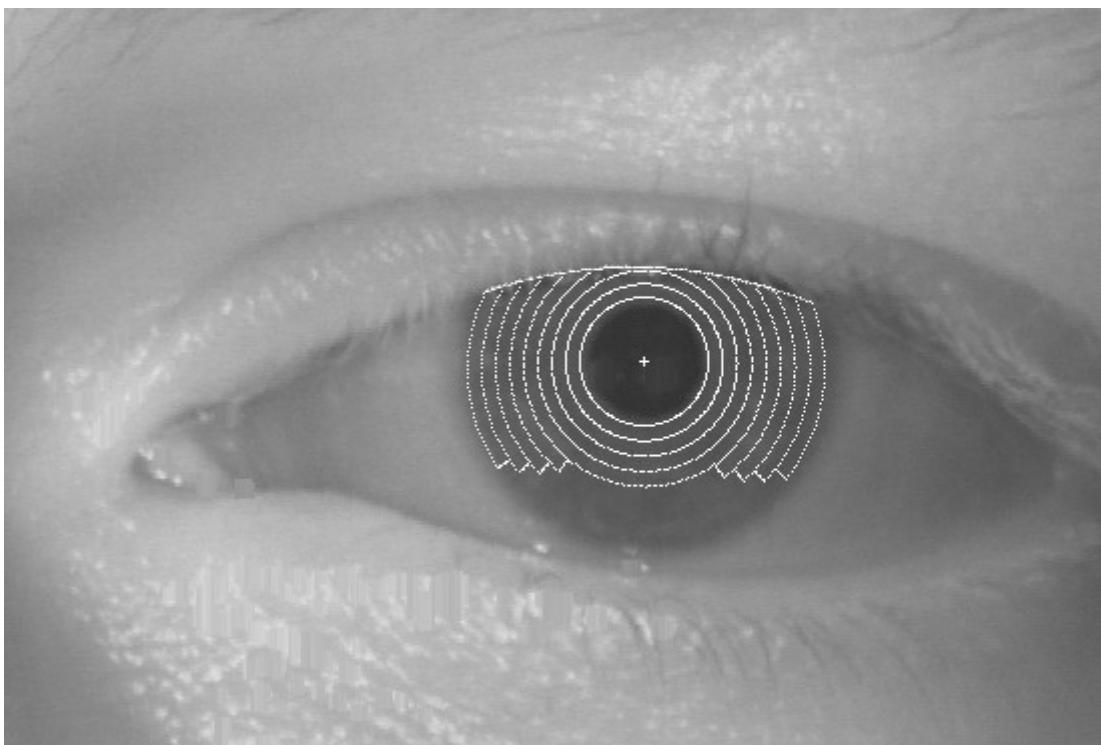


Рисунок 9 определение положение зрачка путем построения концентрических окружностей

Так же, подобная система может с успехом использоваться для запуска хранителя экрана, когда оператор отсутствует, заснул на клавиатуре или просто занят посторонними делами (исследует распечатки, например). Традиционные экранные заставки очень часто ошибаются, определяя неактивность пользователя по отсутствию движений мыши и нажатий на клавиатуру. А если мы смотрим фильм, слайд-шоу или что-нибудь типа того?! Ложные срабатывания крайне раздражают, а увеличение периода неактивности катастрофически снижает их эффективность (кстати, поскольку современные мониторы уже давно не "выгорают", то основная задача хранителей экрана — хладить процессор и автоматически блокировать систему, предотвращая неавторизованный вход посторонних пользователей в случае отлучения оператора на некоторое время).

Для увеличения точности позиционирования рекомендуется использовать лампу, подсвечивающую человеку глаза (естественно, закрытую матовым стеклом и не слепящую) или же LCD-монитор с преобладающим белым фоном на экране. За счет адаптации к свету зрачки сужаются в несколько раз, а благодаря подсветке их контраст с радужкой увеличивается. Как следствие — инструментальная погрешность определения относительного смещения снижается, достигая вполне приемлемого уровня.

Возвращаясь к корпоративным пользователям и системам видео наблюдения, отметим еще одну очень полезную область применения бесконтактных интерфейсов, помогающих оператором выявлять мониторы, обделенные вниманием: если в течении заданного промежутка времени на какой-то конкретный монитор не был направлен ни один взгляд (а на мониторе, судя по обновлению изображения, происходит что-то интересное) раздается звуковой сигнал и вмонтированный в монитор светодиод начинает интенсивно мигать.

Вообще говоря, возможности систем управления взглядом в большей степени ограничены фантазией инженеров, чем самими системами как таковыми. Увы, они опередили свое время и возникли раньше, чем маркетоиды придумали кому их впарить и куда внедрить.

заключение

Бесконтактные интерфейсы стремительно развиваются, захватывая все новые и новые сферы рынка, ранее прочно удерживаемые мышью и клавиатурой. Такого повтора событий не могли представить себе даже фантасты, мечтающие о голосовом управлении, получившего горячую поддержку со стороны производителей сотовых телефонов (действительно, сказать

"таня" или "жена" намного проще, чем выбирать из адресной книги нужный номер), однако, голосовой интерфейс и настольный компьютер — вещи несовместимые.

Во-первых, бесшумные средства ввода (типа мыши) намного более предпочтительны как в офисе, так и дома. Во-вторых, голосовой ввод представляет собой прямую разновидность командной строки, совершенно не вписывающейся в общепринятую оконную концепцию взаимодействия с машиной, в рамках которой намного проще ткнуть мышью в нужный пункт меню, чем озвучить его словами. В-третьих, голосовые системы управления изначально подвержены большим количеством ошибок возникающих как вследствие нечеткого выговора оператора, так и посторонних "помех". Попытки разместить несколько машин с голосовым интерфейсом в одном офисе до сих пор заканчивались неизбежным провалом.

А вот управление взглядом потенциально намного более дееспособно: требуя минимум дополнительных навыков от оператора, оно к тому же очень удачно вписывается в оконную концепцию, позволяя положить обе руки на клавиатуру, забыть про мышь и гонять курсор одними глазами. Не секрет, что метание правой (левой) руки между мышью и клавиатурой не только снижает производительность, но еще и способствует быстрой утомляемости. А синдром "укуса мыши", которым страдают чуть ли не половина пользователей?!

Широкое внедрение систем управления взглядом не только повышает производительность труда, но и позволяет сократить количество производственных травм. В смысле "укусов". Вы уже смеетесь? Напрасно. С юридической точки зрения это в лучшем случае квалифицируется как профессиональное заболевание, лечение которого оплачивается за счет предприятия. Согласен, у нас таких прецедентов не было (и на горизонте обозримой перспективы они навряд ли появятся), но ведь нельзя же все время жить сегодняшним днем. Кому-то нужно и в будущее смотреть. А в том, что будущее за бесконтактными интерфейсами — можно не сомневаться!

>>> врезка управление фокусом в фотоаппаратах

Большинство начинающих фотолюбителей, сменив "мыльницу" на цифровую или пленочную камеру среднего класса, тут же начинают гнать брак, поскольку в отличии от "мыльницы", обладающей большой глубиной резкости, "нормальные" камеры требуют наводки на фокус, которая практически во всех моделях осуществляется автоматически, причем если в моделях пятилетней давности было, как правило, всего три точки фокусировки, то теперь их количество зачастую достигает нескольких десятков, что позволяет встроенному процессору в 90% случаев угадать замысел фотографа, фокусируясь на объекте съемки, даже если он смещен относительно центра кадра.

Проблема однако в том, что в сложных сюжетах (а их, по статистике, не менее 10%), электроника "промахивается" и бьет мимо цели. Быть может, фотограф хотел снять девушку, перекрываемую слегка размытыми ветвями переднего плана. Как объяснить автоматике, что наводить фокус нужно не на ветки, а на девушку, а точнее — на ее глаза, т. к. небольшая размытость остальных частей тела вполне допустима, а нерезкие глаза критики не простят ни за что и никогда.

До сих пор переходило либо переводить аппарат в режим ручной фокусировки (но далеко не на всех моделях он есть), либо отключать "лишние" точки фокусировки, оставив лишь одну (как правило, ту, что в центре), наводить на резкость, делать "lock" (то есть "захват") и перемещать фотоаппарат, компонуя кадровое пространство по своему замыслу. Технически прием простой, но требующий времени. В постановочной или пейзажной фотографии с этим недостатком можно и смирится, поскольку, всегда существует возможность переснять объект по несколько раз, но вот репортажная и спортивная съемка требуют огромной оперативности и счет идет на доли секунды. И вот тут-то управление взглядом очень помогает!

Некоторые производители уже выбросили на рынок модели, где выбор необходимой точки фокусировки осуществляется взглядом (например, Canon EOS-3 с функцией "Eye Controlled Focus"). Куда смотрит фотограф при нажатии кнопки затвора, туда и фокусируется фотоаппарат. Просто, быстро, удобно! Единственный существенный недостаток — достаточно высокая цена, сопоставимая со стоимостью фотоаппарата среднего класса, а потому технология "Eye Controlled Focus" до сих пор остается прерогативой профессиональной технике. Впрочем, как показывает практика, высокие технологии с течением времени проникают и в low-end, так что нам остается только подождать.

Более подробную информацию о принципах действия "Eye Controlled Focus" можно найти на официальном сайте фирмы Canon, прогулявшись по следующей ссылке:

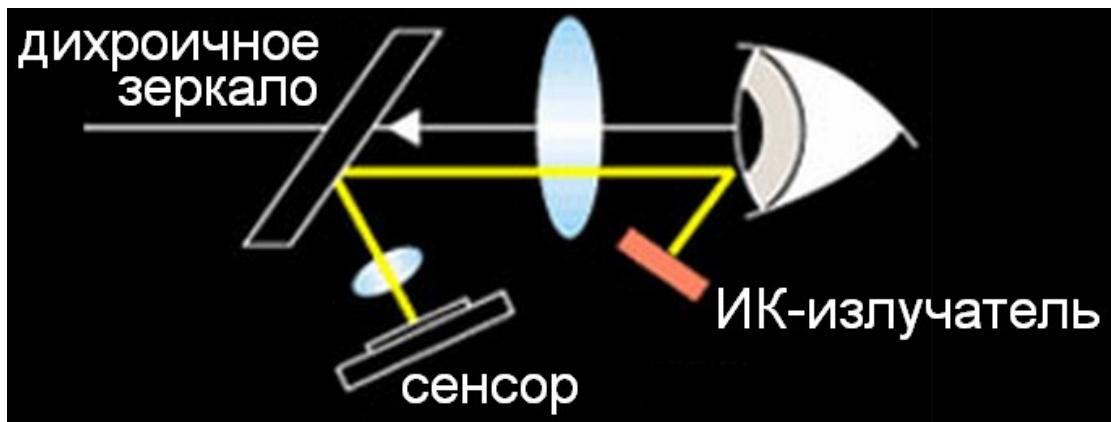


Рисунок 10 система фокусировки, управляемой зрением "Eye Controlled Focus",
применяемая в фотоаппаратах Canon EOS-3