IR Remote Control Theory

The cheapest way to remotely control a device within a visible range is via Infra-Red light. Almost all audio and video equipment can be controlled this way. Due to this wide spread use the required components are quite cheap and readily available, thus making it ideal for us hobbyists to use IR control for our own projects.

This part of my knowledge base will explain the theory of operation of IR remote control and show some details of the protocols that are in use in consumer electronics.

Infra-Red Light

Infra-Red light is actually normal light with a particular colour. We humans can't see this colour because its wave length of about 950nm is below the visible spectrum. That's one of the reasons why IR is chosen for remote control purposes, we want to use it but we're not interested in seeing it. Another reason is because IR LEDs are quite easy to make, and therefore can be very cheap.

Although we humans can't see the Infra-Red light emitted from a remote control doesn't mean we can't make it visible.

A video camera or digital photo camera can "see" the Infra-Red light as you can see in this picture. Even the cheapest cell phones have built in cameras these days. Simply point your remote to such a camera, press any button and you'll see the LED flicker.

Unfortunately for us there are many more sources of Infra-Red light. The sun is the brightest source of all, but there are many others, like: light bulbs, candles, central heating system, and even our body radiates Infra-Red light. In fact everything that radiates heat, also radiates Infra-Red light. Therefore we have to take some precautions to guarantee that our IR message gets across to the receiver without errors.

Modulation

Modulation of the signal on a carrier frequency is the answer to make our signal stand out above the noise. With modulation we make the IR light source blink in a particular frequency. The IR receiver will be tuned to that frequency, so it can ignore everything else.

You can think of this blinking as attracting the receiver's attention. We humans also notice the blinking of yellow lights at construction sites instantly, even in bright daylight.
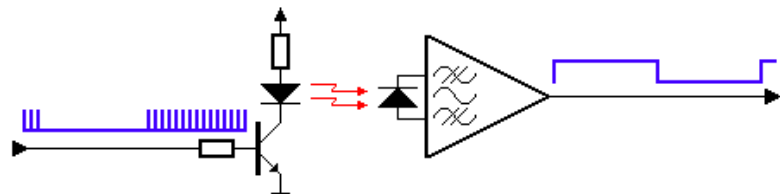
In the picture above you can see a modulated signal driving the IR LED of the transmitter on the left side. The detected signal is coming out of the receiver at the other side.

In serial communication we usually speak of 'marks' and 'spaces'. The 'space' is the default

signal, which is the off state in the transmitter case. No light is emitted during the 'space' state. During the 'mark' state of the signal the IR light is pulsed on and off at a particular frequency. Carrier frequencies between 30kHz and 60kHz are commonly used in consumer electronics. The most common one is 38kHz though.
At the receiver side a 'space' is usually represented by a high level of the receiver's output. A 'mark' is then automatically represented by a low level.

Please note that the 'marks' and 'spaces' are not the 1-s and 0-s we want to transmit. The real relationship between the 'marks' and 'spaces' and the 1-s and 0-s depends on the protocol that's being used. More information about that can be found on the pages that describe the protocols.

The Transmitter

The transmitter usually is a battery powered handset. It should consume as little power as possible, and the IR signal should also be as strong as possible to achieve an acceptable control distance. Preferably it should be shock proof as well.

Many chips are designed to be used as IR transmitters. The older chips were dedicated to only one of the many protocols that were invented. Nowadays very low power micro controllers are used in IR transmitters for the simple reason that they are more flexible in their use. When no button is pressed they are in a very low power sleep mode, in which hardly any current is consumed. The processor wakes up to transmit the appropriate IR command only when a key is pressed.

Quartz crystals are seldom used in such handsets. They are very fragile and tend to break easily when the handset is dropped. Ceramic resonators are much more suitable here, because they can withstand larger physical shocks. The fact that they are a little less accurate is not important.

The current through the LED (or LEDs) can vary from 100mA to well over 1A! In order to get an acceptable control distance the LED currents have to be as high as possible. A trade-off should be made between LED parameters, battery lifetime and maximum control distance. LED currents can be that high because the pulses driving the LEDs are very short. Average power dissipation within the LED should not exceed their maximum value though. You should also see to it that the maximum peek current for the LED is not exceeded. It is quite common to reduce the pulse/pause ratio of the carrier signal to 1/3 or even 1/4. That reduces power requirements or can increase the current through the LED without over heating it to get a longer control distance. All these parameters can be found in the LED's data sheet.
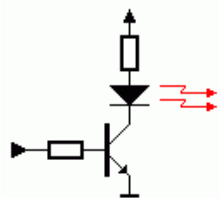
A simple transistor circuit can be used to drive the LED. A transistor with a suitable HFE and switching speed should be selected for this purpose.
The resistor values can simply be calculated using Ohm's law. Remember that the nominal voltage drop over an IR LED is approximately 1.1V.

The normal driver, described above, has one disadvantage. As the battery voltage drops, the current through the LED will decrease as well. This will result in a shorter control distance that can be covered. An emitter follower circuit can avoid this. The 2 diodes in series will limit the pulses on the base of the transistor to 1.2V. The base-emitter voltage of the transistor subtracts 0.6V from that, resulting in a constant amplitude of 0.6V at the emitter when the IR transmitter is

active. This constant amplitude across a constant resistor results in current pulses of a constant magnitude, regardless of the battery voltage. Calculating the current through the LED is simply applying Ohm's law again.
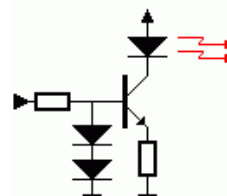
The Receiver

Many different receiver circuits exist on the market. The most important selection criteria are the carrier frequency used and the availability in you region.
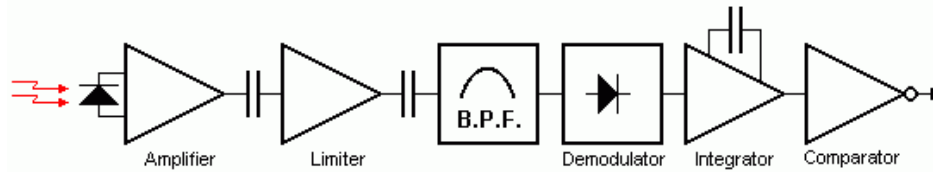
In the picture above you can see a typical block diagram of such an IR receiver. Don't be alarmed if you don't understand this part of the description, for everything is built into one single, easy to use electronic component.
The received IR signal is picked up by the IR detection diode on the left side of the diagram. This signal is amplified and limited by the first 2 stages. The limiter acts as an AGC circuit to get a constant pulse level, regardless of the distance to the handset.
As you can see only the AC signal is sent to the Band Pass Filter. The Band Pass Filter is tuned to the carrier frequency of the handset unit. Common carrier frequencies range from 30kHz to 60kHz in consumer electronics, so it is important to pick the right one.
The next stages are a detector, integrator and comparator. The purpose of these three blocks is to detect the presence of the carrier frequency. If this carrier frequency is present the output of the comparator will be pulled low.

As I said before, all these blocks are integrated into a single electronic component. There are many different manufacturers of these components on the market. And most devices are available in several versions each of which are tuned to a particular carrier frequency.

Please note that the amplifier is set to a very high gain.
Therefore the system tends to start oscillating very easily. Placing a large capacitor of at least 22µF close to the receiver's power connections is mandatory to decouple the power lines. Some data sheets recommend a resistor of 330 Ohms in series with the power supply to further decouple the power supply from the rest of the circuit.

There are several manufacturers of IR receivers on the market. Siemens, Vishay and Telefunken are the main suppliers here in Europe. Siemens has its SFH506-xx series, where xx denotes the carrier frequency of 30, 33, 36, 38, 40 or 56kHz. Telefunken had its TFMS5xx0 and TK18xx series, where xx again indicates the carrier frequency the device is tuned to. It appears that these parts have now become obsolete. They are replaced by the Vishay TSOP12xx, TSOP48xx and TSOP62xx product series.
Sharp, Xiamen Hualian and Japanese Electric are 3 Asian IR receiver producing companies. Sharp has devices with very cryptic ID names, like: GP1UD26xK, GP1UD27xK and GP1UD28xK, where x is related to the carrier frequency. Hualian has it's HRMxx00 series, like the HRM3700 and HRM3800. Japanese Electric has a series of devices that don't include the carrier frequency in the part's ID. The PIC-12042LM is tuned to 36.7kHz, and the PIC12043LM is tuned to 37.9kHz.

The End?

This concludes the theory of operation for IR remote control systems intended for use in consumer electronics. I realize that other ways exist to implement IR control, but I will limit myself to the description above. One of the issues not covered here is security. Security is of no importance if I want to control my VCR or TV set. But when it comes to opening doors or cars it literally becomes a 'key' feature! Maybe I will cover this issue later, but not for now.

I also realize that my small list of manufacturers is far from being complete. It is hardly possible to list every manufacturer here. You can send me an e-mail (../../mailme/index.php) if you have details about other protocols that you feel should be added to my pages.

This page only described the basic theory of operation of IR remote control. It did not describe the protocols that are involved in communication between transmitter and receiver. Many protocols have been designed by different manufacturers. You can find the protocols of some manufacturers navigating the other pages in this chapter using the tools in the right hand top corner of this page.