Sensing and Control in Optical Drives
How to Read Data from a Clear Disc

AMIR H. CHAGHAJERDI

Optical storage devices, such as the compact disc (CD) and digital versatile disc (DVD), are consumer products with control components. In optical storage, data is represented on a plastic disc by the existence or absence of pits, whose reflective characteristics are different from those of the surrounding disc media. Data is read from the disc by reflecting light off the disc to detect these differences. The maximum data transfer rate is directly related to how well the servo systems perform to keep the read/write elements centered over the desired data and at the correct focus.

In CD/DVDs, the pits lie in tracks that form a long spiral from the inner diameter to the outer diameter of the disc. To read the data, the disc is spun, and laser light is focused on the region of the disc containing the track to be read. A sensor monitors the intensity of the reflected light, with a transition from low to high intensity representing the transition from a pit to the nonpit area, called land. Some discs contain two layers at different depths of the disc, increasing the amount of storage. A double-sided disc can contain up to two layers on each side, with four layers total. Because these layers are at different distances below the surface, each can be read by adjusting the focus distance.

TABLE 1 A comparison between the physical characteristics of four optical media, CD, DVD, HD-DVD, and Blu Ray. Since DVD supports double-sided usage, a dual-layer double-sided DVD has four layers with a capacity of 17.1 GB. A 1.2-mm thickness and 120-mm diameter are the standards for CD and DVD. A CD/DVD with 80-mm diameter, with capacity of 192 MB for a CD and 1.4 GB/layer for a DVD, is called mini-CD/DVD.

	CD	DVD	HD-DVD	Blu Ray
Laser Wavelength (nm)	780	640	405	405
Track Pitch (nm)	1600	740	400	320
Capacity/Layer (GB)	0.65	4.7	15	25
Layers/Disc	1	1, 2, 4	1, 2	1, 2
Minimum Pit Length (nm)	834	400	204	149
Operating Speeds (X)	1–52	1–16	1–16	1–16

Digital Object Identifier 10.1109/MCS.2008.920436

Several types of optical storage media are available with different methods of recording data. Some of the characteristics of the discs are listed in Table 1. CD-ROM and DVD-ROM are read only, with the data pits stamped permanently on the disc. CD-R, DVD-R, pronounced "DVD dash R," and DVD+R, pronounced "DVD plus R," can record data only once through a write process that records a pit on the media by melting it. Although there are a few technical differences

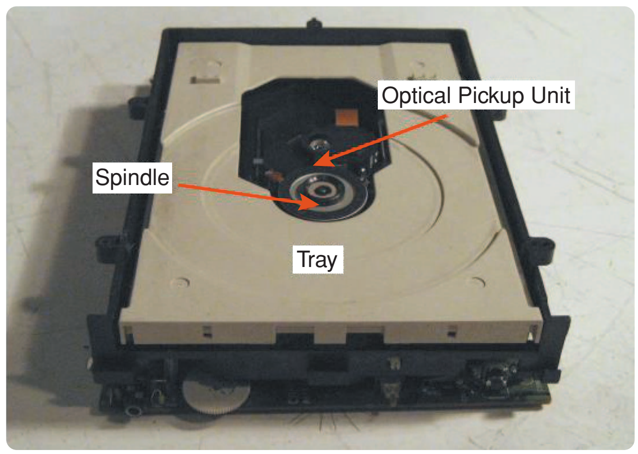

FIGURE 1 CD player with cover removed. The tray carries the CD into the drive, where it is rotated by the spindle. The optical pickup unit is moved to read the information contained on the disc.

between DVD-R and DVD+R, consumers do not see the difference. These formats are developed by two different coalitions of corporations. Today most DVD drives are hybrid and can play or record both formats. Writing requires higher laser power than reading in DVD/CD drives. For example, when operating in read/write speeds of 1–16X, DVD-R requires laser power of 0.4–1 mW for reading versus 10–250 mW for writing.

CD-RW, DVD-RW, and DVD+RW can be rerecorded many times since the reflectivity change due to the write process is reversible, allowing for erasing and rerecording up to around 1000 times. DVD-RAM (random access memory) is also rerecordable, with a rerecording limit of about 100,000 times, which is 100 times higher than DVD-RW and DVD+RW. The DVD-RAM format is different from DVD-RW and DVD+RW, and its structure is close to hard disk drives, where data is stored in concentric tracks.

DVD-RAM media is mainly used in camcorders and personal video recorders.

The read/write lasers, optics, sensors, and some actuators are placed together on an optical pickup unit (OPU), as shown in Figure 1. Feedback control is used to position the OPU over the correct region of the disc as well as to control the focus of the laser used to read and write the data. Deviations in the range of only 2–3% of the distance between tracks, called track pitch (TP), is tolerated before performance is degraded. Table 1 shows that this deviation is equivalent to 32–48 nm for CD and 14.8–22 nm for DVD. Tracking must be maintained in the face of several types of disturbances, classified by the industry as repeatable and nonrepeatable runouts (RROs and NRROs). An example of NRRO is external shock, while disc eccentricity is an example of RRO.

RRO rejection can be achieved through application of a feedforward signal, while NRRO rejection is achieved by a feedback loop. The effect of eccentricity can be considerable in removable media such as DVD/CD since every time the media is loaded, it may be placed at a slightly different center. The result can be a misalignment of up to 200 tracks from one side of the disc to the other. Playability of the optical devices is extensively tested with discs that contain intentional defects representing real-life irregularities, including discs with scratches, eccentricities, fingerprints, black dots, and vertical deviations. These precisely defined reproducible irregularities give manufacturers standard tools for testing the performance and robustness of their products.

In this article, the main servo loops are discussed in detail. These loops include the tracking servo loops (Figure 2) as well as the focus servo loop (Figure 3), which have the highest bandwidth and the most demanding performance requirements.

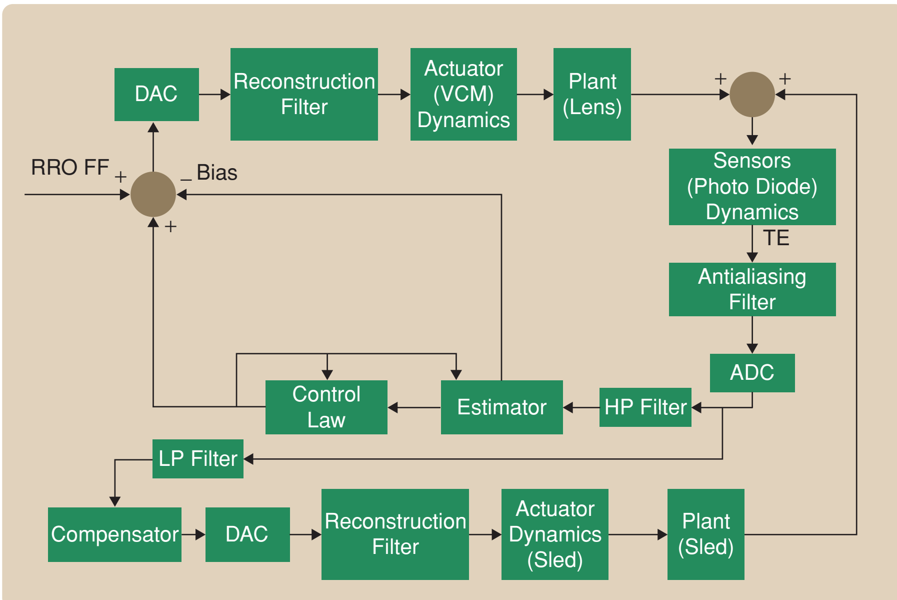

FIGURE 2 The tracking control loops. The lower (and lower bandwidth) loop commands a motor/gear to position the sled that carries the optical pickup unit (OPU). The higher bandwidth loop drives a voicecoil actuator for fine horizontal positioning of the OPU lens. Repeatable runout (RRO) disturbances are estimated separately and enter as a feedforward term. Not shown is the desired track setpoint.

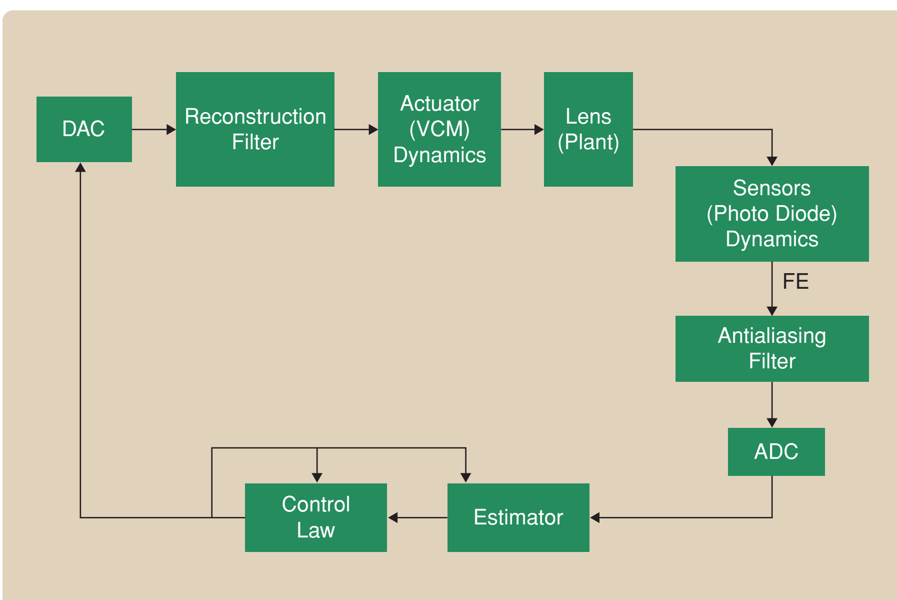

FIGURE 3 Focus control loop. The bandwidth of focus loop is similar to the bandwidth of the fine tracking loop. An estimator is not needed unless a state-space model similar to the fine-tracking loop is used. An analog circuit calculating the focus error is included in the sensor's dynamics. This circuit can calculate only the distance from the focus point and not the absolute position. Hence the setpoint is eliminated from the diagram.

ACTUATION

Optical drives typically have six electromechanical actuators [3], some of which are shown in

The maximum data transfer rate is directly related to how well the servo systems perform to keep the read/write elements centered over the desired data and at the correct focus.

figures 4 and 5. A rotary dc motor, called the spindle motor, is used for spinning the turntable, while a second dc motor is used for loading/unloading (LUL) the disc/tray. The OPU resides on a carrier called a sled, while a third rotary motor, called the sled actuator, is used to move the sled linearly through a worm and gear. A linear voicecoil actuator, called the tracking actuator, is used to move the lens of the OPU horizontally for fine tracking, while a second, called the focus actuator, moves the OPU vertically for focusing. For track following or for short hops within a range of hundreds of tracks, called short seek, the tracking actuator is used, while, for long seeks, the sled actuator performs the positioning. Finally, a sixth actuator, called the tilt actuator, has recently been introduced for changing the tilt angle of the lens. This actuator compensates for manufacturing tolerances and irregularities of the disc surface and lens in order to keep the laser beam perpendicular to the surface. In most cases there is no need for a servo loop for tilt compensation, and startup calibration is good enough.

SENSING

The main goal of the control system is to keep the focus point at the location of the track of interest. A sensor system detects vertical misalignment, called focus error (FE), and horizontal misalignment, called tracking error (TE), using information contained in the reflected laser intensity.

Focus Error

FE is determined by using an astigmatic objective lens and a sensor for the reflected light. The sensor contains four separate photo diodes that divide the sensing area into separate quadrants [3], as shown in Figure 6. When the objective lens is a typical spherical lens, the light reflected from the data pit appears on the detector as a circle, with the radius minimized when the pit is in focus. A large radius indicates that the lens is not focused on the track but does not indicate whether the lens needs to move closer to the focus point. To eliminate the need for more complicated signal processing to determine the correct direction, the objective lens is astigmatic diagonally, giving the lens a different focal point along each diagonal. When the track to be read is between the two focal points, which is the desired read position, the spot on the detector is circular. However,

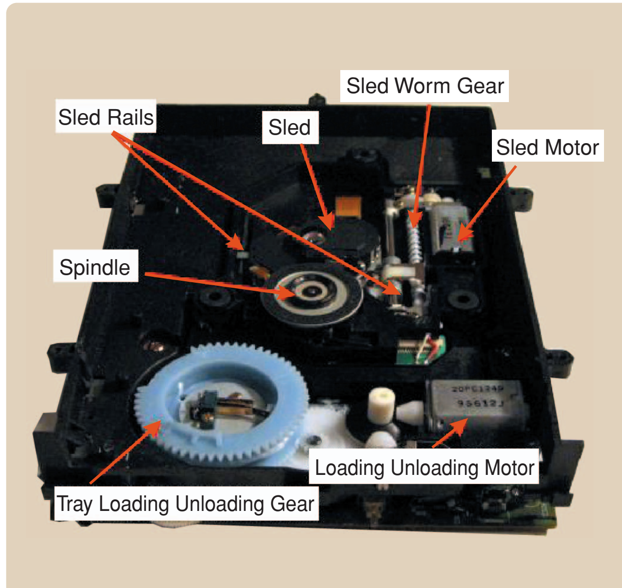

FIGURE 4 CD drive with tray removed. Under the tray some of the actuators are visible, including the spindle, which spins the disc, the loading and unloading mechanism, and the motor and gear that translate the sled containing the optical pickup unit.

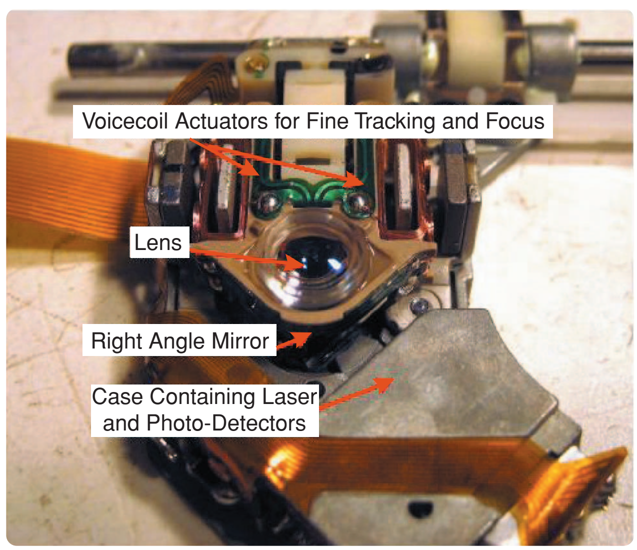

FIGURE 5 The optical pickup unit. The laser light exits the case at the lower right, is directed upward by the right-angle mirror, and is then focused by the lens. The lens is moved horizontally and vertically by voicecoil actuators.

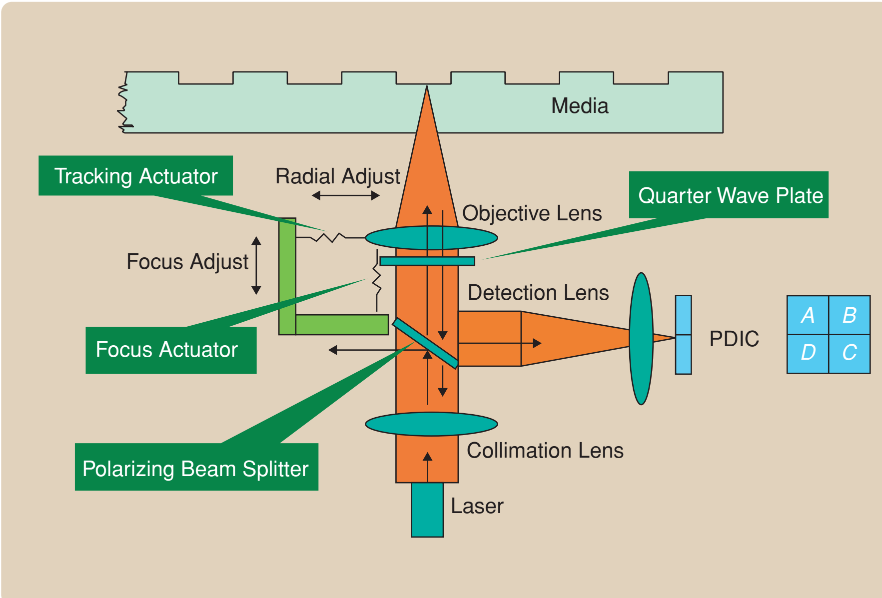

FIGURE 6 The optical pickup unit. One possible arrangement of the optical path is shown here. Laser is shone through collimation lens, polarizing beam splitter, quarter-wave plate, and objective lens before reaching the media. The reflected light from the media goes through objective lens, quarter-wave plate, and detection lens before reaching the photo detectors. A diffraction grating can be also placed between the laser and collimation lens to provide side beams. Side beams follow the same path as the main beam, and are detected on side beam detectors half a track away (not shown). Except for the media, all of the components shown are placed on the optical pickup unit.

when the lens moves away from this position, the lens moves toward one or the other focal point, thus shortening the spot along one axis and lengthening it along the other, which creates an elliptical spot as shown in Figure 7. The FE signal is then calculated from the intensity reaching each quadrant of the detector as

$$FE = A + C - (B + D).$$

As the lens moves up and down, the FE is expected to have the S-curve behavior shown in Figure 7. The error is linear with distance close to the track, within a range of about 10 μm. Beyond this point, the reflected signal becomes diffuse and the FE signal goes back to zero. For details, see [3].

The shape of the S-curve plays an important role at startup. When the disc is first inserted, the servo does not know the position of the data on the disc or the number of data layers. By sweeping through all values of focus, and by examining zero crossings, the number of layers and their focus points can be determined [4].

The FE signal can be supplemented with another signal measure to optimize focus position for the read and write processes. For the read process, maximizing the total signal strength $A + B + C + D$ is relevant, while, for writing, maximizing the push-pull signal of $A + D - (B + C)$ is essential. The location of these maxima versus FE are measured and applied as an offset to the desired FE in the focus loop. This process is sometimes referred to as defocus.

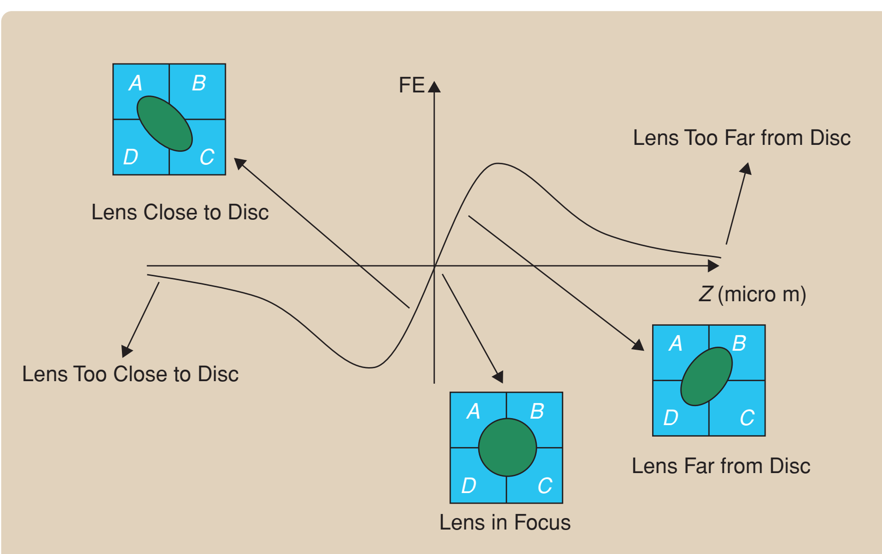

FIGURE 7 Focus error for a typical arrangement of four detectors. Note the characteristic S-curve generated when the lens is moved up and down. The spot is circular at focus and elliptical when away, while the focus-error signal is small in magnitude when the lens is far away from the disc or close to the disc.

Tracking Error

The TE measures the radial distance of the laser spot from the desired track. Figure 8 shows a typical arrangement of data pits in the track grooves, the four data detectors discussed in the previous section, and four supplementary sensors that can also be used for tracking.

When utilizing only the central four detectors, two methods are used to calculate the TE. One method is called differential push-pull (DPP), which does not require that the disc contain data pits but instead relies on the difference in reflective properties between the region where data is supposed to be written, called the groove, and the remainder of the disc, that is, the land. In this method the intensity of light received at the detectors on the left side is compared against the ones on the right side. With four detectors, this comparison is given by

$$DPP = A + D - (B + C).$$

As shown in Figure 8, when the four central detectors are placed exactly on the center of a track, the received intensities of light on the left and right sides are the same, since the laser light is centered on the groove and the light is mainly reflected off the groove. However, when the detectors deviate to the left or right, the laser spot is no longer centered on the groove. As a result, some of the light reflected to the detectors on one side comes from the land, thus causing a difference in the received intensity between the left and right detectors.

The second method for calculating TE, called differential phase detect (DPD), requires data pits on the disc and is not applicable to a blank disc, since blank discs have no pits. The phase difference of the diagonal terms S1 and S2, defined by

$$S1 = A + C,$$
$$S2 = B + D,$$

is used to indicate how well the laser spot is located over the track. When the laser spot is perfectly centered over the track, the signals S1 and S2 plotted versus time are aligned; that is, the phase shift between the signals is zero. However, when the detector deviates to the left or the right, a phase difference between the diagonal terms is present. For example, consider the configuration in Figure 8 with the detectors offset to the right of the track while the track motion is downward. In this case, the B and C detectors do not see as much of the data pits, and the reflected light intensity is fairly steady. On the other hand, detectors A and D receive the majority of light reflected from the pits, while D lags A due to its relative position, giving rise to a phase difference between signals S1 and S2. The magnitude of this delay is proportional to TE. Likewise, when the four detectors deviate to the left, B and C receive the majority of the reflected light, and B leads C. This technique requires that the disc have pits, which is practical for most ROM medias, even though it is recommended by the DVD forum, an international association of DVD manufacturers and users, that DPD be used for TE calculation whenever possible.

Another advantage of DPD is its largely linear response with respect to TE. As shown in Figure 9, sweeping the lens over several tracks results in a piecewise-linear DPD signal. A TE signal that is linear with respect to the physical distance error makes the control loop easier to implement. DPP, however, is not linear with respect to displacement away from the

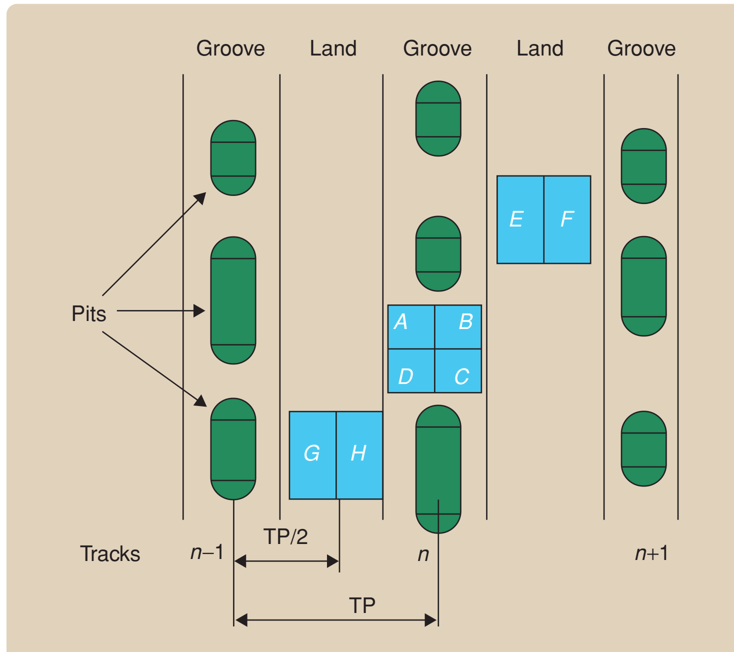

FIGURE 8 Placement of groove, land, pits, and detectors, including main detectors versus side detectors half a track away. The pits are placed at the center of the groove. The side detectors help measure more accurate track-error signals. Each side detector can be a set of two or four photo diodes. The main detectors are placed on the center of the tracks, while the side detectors are placed at the center of the land, which is also called the half-track. This diagram shows three tracks.

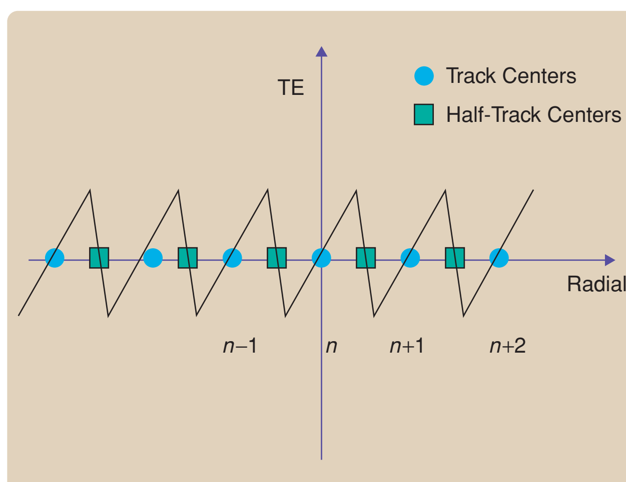

FIGURE 9 Track error, based on differential phase detect, sweeping the lens over tracks. The track-error signal is capable of showing only the distance from the track center or half-track center. Sweeping the lens over tracks generates a track-error signal that can be used to count the number of half tracks.

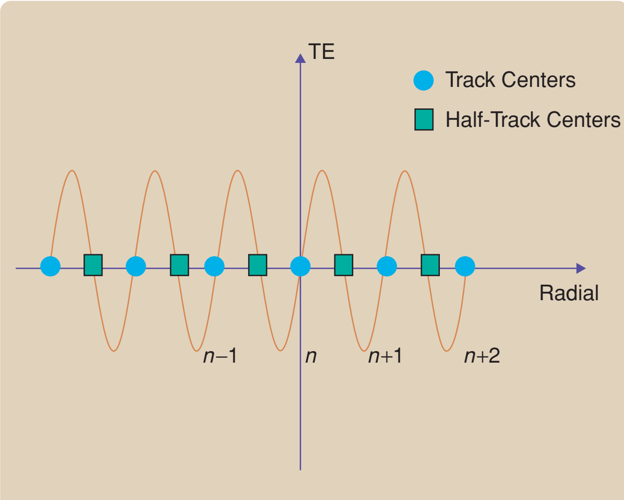

FIGURE 10 Track error, based on differential push pull, sweeping the lens over tracks. The track error calculated in this method has a sinusoidal appearance. Similar to the track error calculated with the differential push-pull method, zero crossings can be used to count the number of half tracks during a seek, which represents the traveled distance.

track or half-track center, and, since sweeping the lens over tracks provides a sinusoidal signal, as shown in Figure 10, some type of linearization is required to ensure a constant loop gain in the tracking servo loop.

The DPP signal may have a dc offset, which is due to displacement of the objective lens. To have a more accurate measure of the TE for DVD, two more sets of detectors, each placed half a track away, can be used [1]. Each extra set can have either two or four detectors. In Figure 8 the additional sets are shown with two detectors. The laser beam is split into a main beam and two sub-beams using a diffraction grating. The main beam is detected by the main or center detectors, while the side beams are detected by the side detectors. The additional detectors provide two push-pull signals, namely, the main beam push-pull (MBPP), which is the same as the DPP signal defined above, and the side beam push-pull (SBPP) defined by

$$\mathrm{SBPP} = E + G - (F + H)\,.$$

When a dc offset is present, the MBPP signal is the sum of the TE and the offset, that is, $\mathrm{MBPP} = \mathrm{TE} + C$, where C is the offset [5]. The SBPP signal contains the same offset but is proportional to TE with the opposite sign, so that $\mathrm{SBPP} = K(-\mathrm{TE} + C)$. Therefore, TE can be constructed with better accuracy using both MBPP and SBPP as

$$\mathrm{TE} = \frac{\mathrm{MBPP} - \mathrm{SBPP}/K}{2}\,,$$

where K is an estimate of the relative intensity of the SBPP signal, calculated during the calibration process [7].

SERVO LOOPS

The six major control loops in optical drives are fine tracking, coarse tracking, focus, spindle, phase-locked loop (PLL), and laser power. Tracking and focus loops have higher bandwidths and are the most complicated loops. The digital servos of both the tracking and focus loops typically include a third- or fourth-order estimator as well as a state-feedback controller. The dynamic behavior of the lens is similar in both the radial and vertical directions and is modeled for control purposes as an isolated mass so that the plant is a double integrator with two states, namely, position and velocity, in a state-space model. A constant force is also included in the model to represent the friction force and tension force from the connecting cable, collectively called the bias force. The bias force varies from track to track. Therefore, a third state is usually added to estimate this bias. Adding the bias term

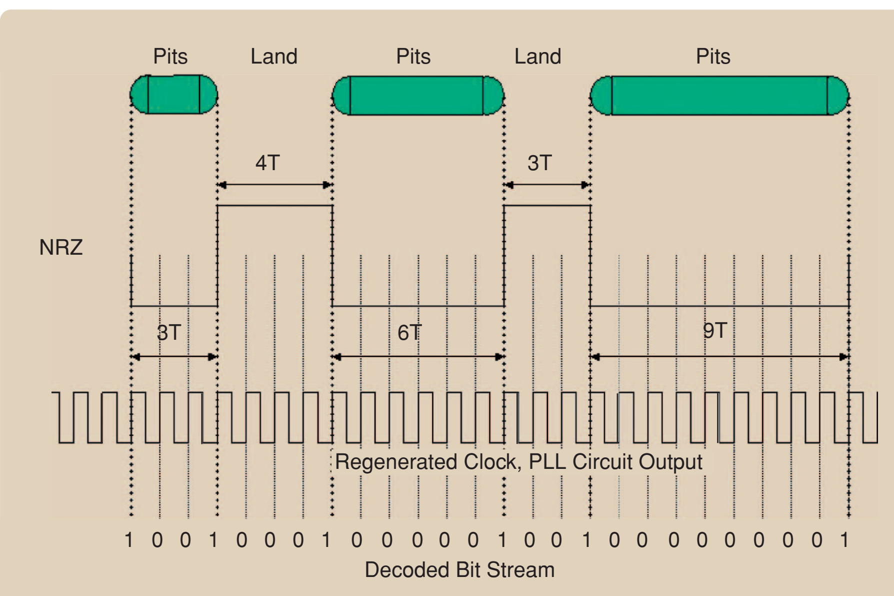

FIGURE 11 Nonreturn-to-zero signal corresponding to actual bits on the media and regenerated clock from a phase-locked-loop (PLL) circuit. To correctly decode the bit stream, the clock frequency needs to be accurate, which motivates the need for a phase-locked-loop circuit.

provides an alternative to appending an integrator to the controller to reduce the steady-state error or runout through increasing the low-frequency gain in the open loop. If the dynamics of the actuator driver are not negligible, then a fourth state, voicecoil current, can be included with the states. If a considerable amount of delay is present in the actuation, then additional states must be included in the discrete-time system model [2].

Although fine tracking (radial lens movement) and coarse tracking (sled movement) are physically coupled, for control design they can be considered decoupled due to the difference in their frequency ranges of operation, with the sled loop having a much lower bandwidth than the fine tracking. The fine-tracking loop uses the TE signal filtered by a highpass filter, while the coarse tracking loop uses a lowpass filter [6]. Figure 2 shows the combined loop consisting of fine tracking and coarse tracking. The focus loop is similar to the fine-tracking loop since the dynamics of the lens are similar in both the radial and vertical directions. Figure 3 shows the focus loop.

The spindle loop is a low-bandwidth loop with a simple controller. Usually a proportional-integrator (PI) controller is sufficient for controlling the spindle. In this loop the setpoint depends on the operating modes. In constant angular velocity (CAV) mode, the setpoint is a constant related to the nominal data rate chosen, for example, 1X or 2X. However, in constant linear velocity (CLV) mode, the setpoint is not only a function of the chosen nominal data rate, but also a function of the radius at which the lens is placed on the disc. Moving from the inside diameter (ID) toward the outside diameter (OD) requires a slower angular velocity setpoint in this mode.

A PLL is used to adjust the sampling rate for decoding the nonreturn-to-zero (NRZ) signal into the actual bits represented on the media. The analog readback signal $A+B+C+D$ is digitized through a slicing method to a NRZ signal [3] with land represented as one and a pit as zero. The binary data are encoded by the pit transitions. A transition between a pit and land or vice versa represents a data one, while lack of a transition represents a data zero. Multiple continuous zeros result in a constant signal, so that, for proper decoding, a clock is needed to set the time corresponding to one bit. Figure 11 shows how this clock is used to decode the stream bits. The clock is synchronized to the NRZ signal using the PLL. Since the spindle speed fluctuates, the accuracy of the clock frequency or period of the clock is vital in decoding. In addition, since the physical length on the medium of a unit pit or land is constant, in CLV mode the PLL-clock frequency is constant from ID to OD. However, in CAV mode this frequency is constantly changing.

Many tradeoffs arise when designing a servo loop for optical drives. The choice of sampling rate is one issue. On the one hand we desire a short sampling time to have smaller phase delay, but we must also ensure enough time to perform all of the calculations for the control signal. The proper servo loop bandwidth is an additional consideration. Increasing the bandwidth of the loop provides faster response and smaller dynamic TE and FE. A high-bandwidth loop, however, is more susceptible to noise and can excite high-frequency modes of the system. In addition, the servo loop must have sufficient phase and gain margin for robust stability. The fact that CD/DVDs operate at different speeds increases the complexity of the servo system design.

CONCLUSIONS

We have described the basic mechanisms for actuation, sensing, and control in optical storage. Key servo loops include tracking, focus, spindle, and the PLL. The speed and quality of the read and write processes depend on how well these servo loops perform.

Recent advances in optical storage technology include the use of blue lasers, whose smaller wavelength allows for detection of smaller pits as well as usage of smaller TP and thus increases the optical drive's capacity. The new needs of the market in consumer products, for example, high-definition content such as HDTV, makes blue ray optical storage likely to become more prevalent in the near future.

ACKNOWLEDGMENTS

I wish to thank Tyrone Vincent for his help in preparing this manuscript.

REFERENCES

[1] S. Yamada, Y. Kuze, K. Watanabe, K. Kondo, and A. Yoshikawa, "Tracking error signal generation device, optical disc apparatus, tracking error Signal generation method and tracking control method," U.S. Patent 5828634, July 10, 2007.

[2] G.F. Franklin, J.D. Powell, and M. Workman, *Digital Control of Dynamic Systems*. Reading, MA: Addison Wesley, 1998.

[3] S.G. Stan, *The CD-ROM Drive, A Brief System Description*. Norwell, MA: Kluwer Academic, Mar. 1998.

[4] I.-W. Hwang, "Method for discriminating a type of a disc and a digital versatile disc," U.S. patent 6061318, May 2000.

[5] K.C. Pohlmann, *The Compact Disc*. London: U.K.: Oxford Univ. Press, 1989.

[6] Y. Zhao, "Non-linear center-error generator for DVD servo control," U.S. Patent 6339565, Jan. 2002.

[7] W. Wu, J. Hsu, and B. Hsu, "Gain calibration device and method for differential push-pull tracking error signals," U.S. patent 7023767, June 2006.

AUTHOR INFORMATION

Amir H. Chaghajerdi (Amir.Chaghajerdi@hitachigst.com) received a Ph.D. in engineering systems in 2002 from the Colorado School of Mines. He is currently with Hitachi Global Storage Technologies in the servo development group at the San Jose enterprise HDD unit.