Coherent transient continuous optical processor

W. Randall Babbitt and John A. Bell

After the absorption profile of an inhomogeneously broadened solid is programmed by two temporally modulated pulses (at least one encoded with a pattern), it can be gated to fix the ground-state spectral population distribution permanently. The subsequent illumination of the solid by an uninterrupted, temporally modulated optical beam results in a coherent transient output signal that represents the correlation of the signal with the stored pattern. Multiple patterns can be stored at different locations on the sample and accessed randomly, enabling fast reprogramming of the processor. A performance analysis of this optical signal processor predicts that real-time continuous processing is possible with a processor bandwidth that exceeds 5 GHz, a time-bandwidth product that exceeds $10^4$, and a pattern storage density that exceeds $10^5$ patterns per square centimeter.

Key words: Optical coherent transients, photon echoes, spectral hole burning, optical signal processing, optical correlator.

1. Introduction
The continuous optical correlator presented here is based on the phenomena of coherent transients, also referred to as photon echoes or time-domain spectral hole burning. Long-term coherent transient storage of temporally modulated optical data pulses in cryogenic solids has been demonstrated by several authors. The temporally encoded information is stored as a spectral interference pattern in the ground- or excited-state inhomogeneously broadened population distribution. Permanent storage and nondestructive reading capabilities can be achieved by a gating process that permanently fixes the spectral population grating. Such processes have been demonstrated for frequency-domain spectral hole burning. Optical coherent transient techniques can also be used to perform convolutions and correlations of three temporally modulated light pulses. The coherent transient output signal represents the cross correlation of the first pulse’s temporal waveform with the convolution of the second and the third pulses’ waveforms. The processor responds to the electric-field amplitudes of the input pulses and thus can fully process amplitude-, phase-, and frequency-modulated signals. In previously proposed implementations, it has been asserted that to obtain high-fidelity correlations, the durations of the modulated input pulses must be less than the homogeneous decay time of the absorbers. To search an uninterrupted data stream (data streams longer than the homogeneous decay time) for a given pattern requires that the signal be broken up into overlapping segments that must be processed separately. To process each segment, the optical pattern/reference pulses would have to be regenerated and introduced into the medium. Thus three modulated light beams with appropriate delays need to be generated for each segment processed. The resultant signals would then have to be additionally processed to obtain the true correlation. These limitations greatly increase the processor’s latency time and make real-time processing of continuous signals impossible.

In this paper we introduce a coherent transient signal-processing technique that enables the real-time processing of continuous input signals after a single programming stage. The unique features of the continuous optical correlator and the techniques for programming it and obtaining high-fidelity correlation peaks are discussed. Then a more detailed analysis of the continuous processor’s potential performance is presented. Data rates that exceed 5 GHz, time-bandwidth products that exceed $10^4$, and pattern storage densities that exceed $10^5$ patterns per square centimeter are predicted.
Coherent transient triple-product correlator: Consider three temporally encoded light pulses. These light pulses may be amplitude, phase, or frequency encoded. When an inhomogeneously broadened two-level absorber is reasonantly excited by these pulses in sequence, the resultant coherent transient output signal \( E_s(t) \) is given by

\[
E_s(t) \propto \int_{-\infty}^{\infty} dt'' E_1(t'') \int_{-\infty}^{\infty} dt' E_2(t') E_3(t'' + t - t'), \tag{1}
\]

where \( E_j(t) \) represent the electric field of the \( j \)th pulse. The output signal represents the convolution of the first data pulse with the convolution of the second and third data pulses. Either the convolution or correlation of only two pulses can be achieved if one of the three data pulses has only a single subpulse whose duration is less than the shortest temporal feature of the other two pulses. Expression (1) for the output is valid regardless of the modulation characteristics of the data pulses. Thus a coherent transient correlator is capable of simultaneously performing phase and frequency correlations, as well as amplitude correlations. The data bandwidth of the triple product correlator is ultimately limited by the inhomogeneous bandwidth of the absorbing transition. For such transitions in solids, data bandwidths can range from a gigahertz to a terahertz. The correlate's time–bandwidth product is limited by the ratio of the inhomogeneous to the homogeneous broadening, which was measured to be as high as \( 10^7 \) in one solid.\(^1\)

2. Coherent Transient Continuous Optical Processor

The assertion made above that all the pulses must be less than the homogeneous decay time is not valid in all cases. Although the time–bandwidth product is limited by the homogeneous dephasing time, the data pulse lengths are not. The duration of the third pulse is not even limited by the population decay time and can actually be infinite. To accomplish this improvement, we must add a step to the programming of the coherent transient correlator that permanently stores the ground-state spectral population grating produced by the first two data pulses. This is referred to as the gating step. The programming, gating, and processing steps of the coherent transient continuous optical correlator are illustrated in Fig. 1.

![Fig. 1](image)

*Fig. 1. Three steps for signal processing: programming the inhomogeneously broadened solid (IBS), gating the IBS, and processing the continuous input signal. As drawn, the output signal propagates in the same direction as the reference pulse.*

A. Programming and Gating the Continuous Processor

Programming the continuous processor is accomplished by illuminating the material with two resonant light pulses: a pattern pulse and a reference pulse. The effect of these first two pulses is identical to the first two pulses in the triple product correlator in that the resultant frequency-dependent population gratings in ground and excited states of the two-level absorbers have components that are proportional to the Fourier transforms of pulses one and two.\(^1\) Immediately after the material is programmed, the material is gated. The gating process permanently alters a significant portion of the excited absorbers such that they are no longer resonant with the modulated light pulses. The timing of the programming and the gating pulses is illustrated in Fig. 2(a). After the programming and gating process, the spectral distribution of the absorbers that were unaltered by the gating process, \( N'(\Omega) \), is proportional to the spectral interference of pulses one and two. Assuming the first pulse is a pattern pulse and the second pulse is a temporally brief reference pulse,

\[
N'(\Omega) \propto E_1^*(\Omega) \exp(-i\Omega t_2) + c.c. + \cdots, \tag{2}
\]

where \( E_1(\Omega) \) is the Fourier transform of pulse one, \( t_2 \) is the time at which the brief second pulse passes through the material,\(^2\) and c.c. indicates the complex conjugate of the previous term. This storage of the spectral interference of the two modulated light pulses in a ground-state population grating is analogous to storage of the spatial interference of two images in intensity gratings on holographic film.

The gating step can be accomplished by illuminating the material with an optical pulse resonant with a
transition from the excited state to another state that
decays into a metastable state whose energy is signifi-
cantly different from the initial ground state. The
gating photons could be of sufficient energy to pho-
toionize the excited absorbers and, thus, permanently
alter their energy level structure. The electron from
the photoionized absorbers could be trapped by an
acceptor ion present in the material. Such a process
has been demonstrated for holes burned into an
inhomogeneous line. Other possible gating meth-
ods include two-step photodissociation and two-step
donor-acceptor electron transfer. If the gating pro-
cess is inefficient, it can be repeated until a sufficient
ground-state population grating is achieved.

For applications that require temporary processing
for a time duration that is much longer than the
upper-state lifetime but short compared with the
ground-state storage time, it may not be necessary to
gate the absorbers. A single photon storage process
could be used. Materials that have multiple ground
state (i.e., hyperfine split ground states) can store
population gratings by having a percentage of the
excited state absorbers decay to a level other than the
original ground state. In ground states with nuclear
hyperfine splittings, these population gratings can
persist for days at liquid helium temperatures. These
materials have two disadvantages. The first is
that the processing stage is destructive. The third
laser pulse continually excites the ground-state at-
oms, which can then decay to another state, depleting
the stored population grating. The available process-
ing time is thus limited by the branching ratio of the
excited state and by the excitation rate required for
obtaining a satisfactory signal-to-noise ratio. The
second disadvantage is that the ground-state splitting
may be smaller than the width of the inhomogeneous
line. For example, nuclear hyperfine splittings are
of the order of tens of megahertz, and inhomogeneous
linewidths are of the order of several gigahertz.
Level overlap will lead to coherent beating and unde-
sired modulation of the output signal. The data
bandwidth would be severely limited by the smallest
ground-state splitting. In this paper, only gated
coherent transient processors are analyzed.

B. Processing Stage
In the processing stage, the material is probed by an
uninterrupted, temporally modulated input signal
that is resonant with the unaltered two-level absorbers.
This probing is nondestructive because, in the
absence of the gating field, any unaltered absorber
that is excited by the input signal will decay to its
initial ground state. Provided the continuous input
data signal $E_d(t)$ does not saturate the absorbers, the
time-dependent frequency components of the polariza-
tion created by the input signal are given by

$$P(t, \Omega) \propto \exp(-i\Omega t)N'(-\Omega) \int_{-\infty}^{\infty} dt' E_d(t') \exp(-i\Omega t').$$  (3)

The resultant output signal is proportional to the
time-dependent polarization and is given by

$$E_o(t) \propto \int_{-\infty}^{\infty} d\Omega P(t, \Omega) \int_{-\infty}^{\infty} dt' E_d(t' + t - t_0) E_1(t' - t_1),$$  (4)

where $t_1$ marks the start of the pulse one, and $t_0 = t_2 - t_1$. The output signal represents the delayed
continuous convolution of the input signal with the
temporal waveform of the first pulse, assuming the
first pulse ends before the second pulse starts (see Fig. 2(b)).

There are no restrictions on the duration of the
input signal, provided all the excited absorbers relax
to the previous ground states. (Although the input
signal may have an indefinite duration with an indefi-
nite number of breaks, it is often referred to as the
third pulse throughout the remainder of this paper.)
Multiphoton processes may eventually lead to degra-
dation of the stored information and thus impose
limits on the third input pulse. To achieve high-
fidelity output signals, we must reduce the intensity
of the third pulse to a level at which saturation effects
are negligible. This is discussed in detail below.
As the processor can accept data continuously, the
latency time of the processor is effectively zero, and
the propagation delay through the processor is roughly
just the sum of the durations of the two programming
pulses.

The output direction is $k_c = k_3 + k_2 - k_1$, where $k_1$, $k_2$, $k_3$, and $k_c$ are the wave vectors for the first,
second, and third pulses and the output signal,
respectively. The output is optimized when the phase
matching condition, $|k_c| = |k_3|$, is satisfied. As the
material is preprogrammed by pulses 1 and 2, during
the processing state the output signal need only be
isolated from the third pulse. This can be accom-
plished by introducing slight angular separation be-
 tween the two programming pulses (as illustrated in
Fig. 1) or by having the pulses be counterpropagating.
Large angles are avoided because of beam-overlap
considerations, not phase-matching constraints.

C. Reprogramming the Processor
There are two potential methods for reprogramming
the processor. The old pattern information could be
erased by a reversible gating process and a new
pattern programmed into the same spatial volume.
In the second method, a two-dimensional spatial
array of patterns is stored in the material, and
different patterns are randomly accessed spatially.
Access times of a few microseconds could be achieved
with acousto-optic or electro-optic deflectors. The
pattern storage densities are calculated below and can
exceed $10^9$ patterns per square centimeter. There-
fore, in most cases there would be no need to intro-
duce the two new programming pulses (pulses 1 and 2).
This greatly reduces the complexity of the pro-
cessing stage and reduces the latency time during

1540 APPLIED OPTICS / Vol. 33, No. 8 / 10 March 1994
reprogramming as the material acts basically as an array of passive filters.

3. Analysis of the Processor’s Performance

The following is an analysis of the performance of a continuous optical processor for the case in which the medium is gated after the second pulse. The effects of homogeneous and inhomogeneous broadening, coherent and incoherent saturation, output efficiency, shot noise, spatial cross talk, diffraction, and local heating are considered in evaluating the processor’s performance. A method for obtaining optimal performance is prescribed, and an example is given. The determination of the maximum achievable time–bandwidth products, processing bandwidths, and storage densities requires the following input parameters: the wavelength, Einstein coefficient, inhomogeneous bandwidth, and homogeneous bandwidth of the absorbing transition; the absorption length and index of refraction of the material; the maximum permissible levels of distortions and cross talk; the gating and detection efficiency; and the desired number of photoelectrons generated by the peak of the output signal. Determining the performance also requires that limits be placed on the intensities of the input signal and the pattern pulse and on the characteristics of their power spectrums. These parameters and limits are defined and discussed below.

A. Assumptions

For convenience, assume (1) the inhomogeneously broadened absorber has a uniform spectral density of absorbers over the inhomogeneous bandwidth, \( \Delta \nu_i \), which is centered around \( \nu = c/\lambda \), where \( \lambda \) is the wavelength of the light pulses’ optical carrier; (2) the inhomogeneous bandwidth is much greater than the homogeneous bandwidth, \( \Delta \nu_h \); (3) the spectrum of the uninterrupted data stream has a bandwidth less than or equal to \( \Delta \nu_d \), which is defined as the processor’s data bandwidth; (4) the data stream has the limitation such that, averaged over any given time interval, its dynamic power spectrum has the characteristic that the maximum power of any frequency component is less than some specified constant times the average power of the data stream; and (5) the intensity of the data stream averaged over a time interval equal to the dephasing time of the absorbing transition is fairly constant. Making these assumptions greatly simplifies the analysis while introducing only a small error with respect to a model that takes into account the true shape of the inhomogeneous line and the data stream’s spectrum. When the specific material and data stream properties are known, the required modifications are straightforward.

The absorption coefficient for a flat inhomogeneously broadened line with an inhomogeneous bandwidth, \( \Delta \nu_i \), is given by

\[
\alpha = \frac{N \lambda^2 \sigma_{10}}{8 \pi n^2 \Delta \nu_i},
\]

where \( N \) is the concentration of absorbers in the material, \( \sigma_{10} \) is the Einstein coefficient for the transition from the ground (level 0) to excited (level 1) state, \( \lambda \) is the vacuum wavelength of the transition, and \( n \) is the index of refraction of the host material.

B. Coherent Saturation

Assume that the medium responds to first order as a nonmagnetic, linear medium. The intensity \( I_d(t) \) and electric-field real amplitude \( E_d(t) \) of the continuous third data pulse are related by

\[
I_d(t) = \frac{cn(|E_d(t)|^2)_\nu}{4\pi},
\]

where \( c \) is the speed of light in a vacuum, and \( (\cdot)_\nu \) represents the average of the enclosed function over a time interval \( \delta t \). The time interval \( \delta t \) is long with respect to an optical period but short with respect to the reciprocal of the data bandwidth. The effects of coherent saturation can be evaluated by calculating the extent to which the absorbers are coherently driven in a time period that is equal to the excitation transition’s homogeneous dephasing time, \( T_2 = 1/(\pi \Delta \nu_h) \). The Fourier components of the electric-field amplitude at time \( t \) taken over a time interval \( T_2 \) are given by

\[
\tilde{E}_d(v) = \int_{t-T_2/2}^{t+T_2/2} dt' E_d(t') \exp(-i2\pi vt'),
\]

where \( \tilde{E}_d(v) \) is double peaked, with each peak grouped about the positive and the negative optical carrier frequencies (\( \pm c/\lambda \)), respectively. It should be noted that as there is no assumption that \( E_d(t) \) is repetitive, \( \tilde{E}_d(v) \) is a time-varying quantity and is thus referred to as a dynamic Fourier component. Because \( E_d(t) \) is real, \( \tilde{E}_d(-v) = \tilde{E}_d^*(v) \). In accordance with our fourth assumption in Subsection 3.A, we define \( \gamma_c \) as the ratio of the maximum allowable magnitude squared of the dynamic Fourier components to the average magnitude squared of the Fourier components, such that at any given time \( t \),

\[
\gamma_c \geq \frac{\max[|\tilde{E}_d(v)|^2]}{(\langle |\tilde{E}_d(v)|^2 \rangle)_{\Delta \nu_d}},
\]

where \( \langle \cdot \rangle_{\Delta \nu_d} \) represents the frequency average over \( \Delta \nu_d \). The average is taken only around the positive carrier frequency. Let \( I_d \) be defined as the maximum intensity of the data stream when averaged over any time interval \( T_2 \). Relation (8) and \( I_d \) are limits of the characteristics of the data signal that must be obtained to perform the predicted calculations. Assuming that the power in the data stream outside the data bandwidth is negligible, Parseval’s theorem yields

\[
\langle |\tilde{E}_d(v)|^2 \rangle_{\Delta \nu_d} = \frac{T_2}{2\Delta \nu_d} \langle |E_d(t)|^2 \rangle_{T_2} \leq \frac{2\pi T_2 I_d}{nc\Delta \nu_d}.
\]

The maximum pulse area at time \( t \) at a given
frequency over the time interval $T_2$ is given by\(^1\)

\[
\theta_{\text{max}}(t) = \frac{4\pi p \text{max}[|\vec{E}_d(v)|]}{h} = \frac{4\pi p}{h} \left(\frac{2\pi \gamma_c T_2 \bar{I}_d}{nc \Delta \nu_d}\right)^{1/2} \leq \theta_t,
\]

where $\theta_t$ is defined as the maximum data pulse area seen at any frequency and at any time over a time interval $T_2$,

\[
p = \frac{(3.\sqrt{10}h \lambda^3)^{1/2}}{8\pi^2}
\]

is the dipole moment, and $h$ is Planck's constant. It is assumed that the unit vectors of the dipole moment and electric field are aligned. The maximum allowable photon flux of the continuous third data pulse, $F_c$, is

\[
F_c = \frac{\lambda \bar{I}_d}{hc} \leq \frac{2\pi\theta_t^2 n \Delta \nu_d}{3\gamma_c \sqrt{10} T_2 \lambda^2}.
\]

By establishing an acceptable level of nonlinearity in the medium's response to the electric-field amplitudes, we can set the value of the maximum allowable pulse area. At the onset of coherent saturation, the medium's response to the peak components in the pulse area. At the onset of coherent saturation, the medium's response to the electric-field amplitudes, we can set the value of the maximum allowable pulse area. At the onset of coherent saturation, the medium's response to the electric-field amplitude over the time interval.

For example, if 10% nonlinearity is acceptable, the maximum allowable pulse area is approximately $\pi/4$.

C. Incoherent Saturation

As the duration of the third pulse can be much longer than the homogeneous decay time of the excitation transition, incoherent saturation of the transition must be considered. The coherent transient output signal's electric-field amplitude is proportional to the population difference between the ground state (level 0) and the excited state (level 1). Take the general case in which the decay from level 1 to level 0 has a bottleneck, which is labeled level 2. Assumption (5) above assumes that the dynamic power spectrum of the continuous third pulse, when averaged over the effective decay time of upper-state population, is roughly uniform over the data bandwidth. The effective decay time of the upper stage, $\tau'$, is discussed below. Define $\gamma_i$ to be the maximum allowable ratio of the peak value of the time-averaged power spectrum to the average value of the time-averaged power spectrum. More precisely,

\[
\gamma_i \geq \frac{\text{max}[(|\vec{E}_d(v)|^2)_{\tau}]}{\langle (|\vec{E}_d(v)|^2)_{\tau} \rangle_{\nu_d}}
\]

at any given time $t$. This is a further limit on the characteristics of the input data signal. The time-varying induced transition rate $\Gamma$ between levels 0 and 1 at the peak of the power spectrum is thus governed by

\[
R \leq \frac{\gamma_i n \Delta \nu_d F_d}{N \Delta \nu_d} \equiv R_{\text{max}},
\]

where $F_d$ is the incident photon flux of the continuous third pulse averaged over $\tau'$. If the effective population decay time is sufficiently greater than the coherence decay time $T_2$ and the peaks in the power spectrum are randomly distributed, then $\gamma_i = 1$. If the effective decay time is comparable with the coherence decay time or if there are peaks in the power spectrum that do not average out over time, then $\gamma_i$ can be as large as $\gamma_c$.

Define $\tau_{10}, \tau_{12}$, and $\tau_{20}$ as the decay times from levels 1 to 0, 1 to 2, and 2 to 0, respectively. When the rate equations are solved, the maximum population difference is found to be proportional to

\[
\frac{1}{1 + 2R_{\text{max}}\tau'},
\]

where $\tau'$ is the effective decay time of level 1 and is given by

\[
\tau' = \frac{\tau_{10} (2\tau_{20} + \tau_{12})}{\tau_{12} + \tau_{10}}.
\]

The values of $\tau'$ under various limits are

\[
\begin{align*}
\text{for } & \tau_{10} \ll \tau_{12}, \tau_{20} \ll \tau_{12}: \quad \tau' \approx \tau_{10}, \\
\text{for } & \tau_{10} \ll \tau_{12}, \tau_{20} \gg \tau_{12}: \quad \tau' \approx (\tau_{10} \tau_{20})/(2\tau_{12}), \\
\text{for } & \tau_{10} \gg \tau_{12}, \tau_{20} \ll \tau_{12}: \quad \tau' \approx \tau_{12}, \\
\text{for } & \tau_{10} \gg \tau_{12}, \tau_{20} \gg \tau_{12}: \quad \tau' \approx \tau_{20}/2.
\end{align*}
\]

In Subsection 3.B on coherent saturation, an acceptable level of nonlinearity for the processor, $\epsilon$, was introduced. The nonlinearity that is due to incoherent saturation is roughly equal to the deviation of expression (16) from unity or approximately $2R' \tau'$ for $2R' \tau' \ll 1$. In order that incoherent saturation does not introduce nonlinearities greater than $\epsilon$, the condition $2R' \tau' \leq \epsilon$ must hold. Thus the maximum photon flux of the continuous third pulse that is limited by incoherent saturation is given by

\[
F_i = \frac{4\pi n^2 \Delta \nu_d \epsilon}{\lambda^2 \sqrt{10} \tau' \gamma_i}.
\]

Here (and throughout the paper) it is assumed that the nonlinearities that are due to different effects are...
uncorrelated, although further study is needed to determine the extent to which this is true. To minimize both coherent and incoherent saturation, the maximum allowable photon flux of the continuous third pulse must be less than or equal to the minimum of $F_c$ in Eq. (12) and $F_i$ in Eq. (22). In accordance with assumption (5), the time average of the intensity over an interval $T_2$ is roughly equal to the time average over an interval $\tau'$. If the coherence decay time were a fixed property of the material but the population decay time were variable (i.e., by optical pumping of the bottleneck state to a short-lived state), the optimal effective decay time is obtained by setting $F_d = F_i = F_c$, which yields

$$\tau' = \left(\frac{6\pi\gamma e L}{\gamma_0 t^2}\right) T_2.$$  

(23)

Shorter decay times lead to less saturation, but may lower gating efficiencies because in the programming stage there is less time to gate the excited absorbers.

D. Output Signal Efficiency

Define the peak output efficiency $\eta_{peak}$ as the ratio of the intensity of the output autocorrelation peaks to the average intensity of the continuous third pulse. The dependence of $\eta_{peak}$ on (1) the gating efficiency, (2) the properties of the stored pattern pulses, (3) the homogeneous decay time, (4) the gain efficiency of the material, and (5) the spatial profile of the input beams are discussed below.

Gating Efficiency

Define the gating efficiency $\eta_{gate}$ as the ratio of the number of absorbers that are permanently altered (i.e., photoionized) by the gating pulse to the number of absorbers that were in the excited state just before the gating pulse (for accumulated population gratings, replace $\eta_{gate}$ with an appropriate accumulated grating efficiency). The output signal electric-field amplitude is proportional to the population difference at the time of recall and thus is proportional to $\eta_{gate}$. The maximum obtainable electric field of a gated coherent transient (when $\eta_{gate} = 1$) is down by a factor of 2 from that of a nongated coherent transient because of the elimination of the excited-state population grating that contains half of the information after the second pulse. The output signal’s electric field is equal to $\eta_{gate}/2$ times that of a nongated coherent transient.

Pattern Pulse Efficiency

To calculate the intensity of the output signal, we must choose a particular processor configuration. The case considered below is one in which a single pattern pulse is programmed during the programming stage. Assume that the second pulse consists of a single subpulse with a pulse area equal to $\theta_2$. Let the first pulse be a pattern pulse of total duration $\tau_{pat}$. The electric field of the output signal in the limit of an optically thin sample ($\alpha L \ll 1$) is given by

$$E_d(t) = \eta_{gate}\alpha L \sin(\theta_2) \frac{\pi \tau_{pat}}{h} \int_{-\tau_{pat}/2}^{+\tau_{pat}/2} d\tau E_{pat}(\tau) E_d(t + \tau - \tau_{pat}) \times \exp[-2(\tau_{21} - \tau)/T_2].$$  

(24)

where $L$ is the interaction length in the material, $E_{pat}(t)$ is the real electric field of the pattern (first) pulse, and $\tau_{21}$ is the temporal separation between pulses one and two. The time origin $t = 0$ is defined such that $E_{pat}(t)$ is nonzero only from $t = -\tau_{pat}/2$ to $t = +\tau_{pat}/2$. Assuming that pulse two immediately follows the pattern pulse, then $\tau_{21} = \tau_{pat}/2$. The third pulse $E_d(t)$ can start at any time after the material has been programmed and gated. The exponential term in the integral takes into account the effects of homogeneous coherence decay.

Assume that, over a time interval $\tau_{pat}$ about time $t_p$, the third pulse matches the first pattern pulse, i.e.,

$$E_d(t_p + t') = \left[\frac{I_d(t_p + t')}{I_d(t')}\right]^{1/2} E_{pat}(t') \approx \left[\frac{\bar{I}_d}{I_{pat}}\right]^{1/2} E_{pat}(t'),$$  

(25)

where it is assumed that the intensity of the third pulse when averaged over $\tau_{pat}$ is roughly equal to $\bar{I}_d$ and $I_{pat}$ is the average intensity of the first pulse. This assumption is valid if (1) $\tau_{pat}$ is not significantly less than $T_2$, (2) $\tau_{pat}$ is much greater than $1/\Delta \nu_d$, and (3) the average intensity of the third pulse is relatively constant. The maximum allowable intensity of the first pulse $(I_{pat})^{1/2} \tau_{pat}$ is governed by coherent saturation and is equivalent to the expression in Eq. (12) if $T_2$ is replaced with $\tau_{pat}$ and $\gamma_c$ is replaced with $\gamma_{pat}$. The new parameter $\gamma_{pat}$ is introduced to allow for the case in which the spectral characteristics of the pattern pulse differ significantly from those of the continuous pulse three. If the exponential decay term is temporarily ignored, the intensity of the output signal without homogeneous decay at the correlation peak is

$$\lim_{T_2 \to \infty} \left[I_d(t_p)\right] \approx \eta_{gate}^2\alpha L^2 \sin^2(\theta_2) \frac{\theta_2 \tau_{pat} \Delta \nu_d}{16 \gamma_{c,pat}^2 I_d}. $$  

(26)

Note that the intensity of the output peak is proportional to the product $\tau_{pat} \Delta \nu_d$, the time–bandwidth product of the processor.

Homogeneous Decay Efficiency

The effect of homogeneous decay on the correlator’s operation is to put less weight on the start of the pattern than on the end of it and to reduce the output efficiency. The maximum field nonlinearity that is due to the homogeneous decay is $[1 - \exp(-2\tau_{pat}/T_2)]^{1/2}$. By bounding this value with the acceptable nonlinearity $\epsilon$, a maximum value for $\tau_{pat}$ can be obtained.

10 March 1994 / Vol. 33, No. 8 / APPLIED OPTICS 1543
ratio of $\tau_{\text{pat}}/T_2$ is
\[
\beta_{\text{pat}} = (\tau_{\text{pat}}/T_2) \leq (\epsilon/2).
\] (27)

To estimate the reduction in the autocorrelation peak intensity that is due to homogeneous decay, let $|E_{\text{pat}}(t)|$ be roughly constant (as in the case of phase-modulated input pulses), so that the electric fields can be taken out of the integral in Eq. (24). The resultant efficiency that is due to homogeneous decay is
\[
\eta_{\text{decay}} = \lim_{\tau_{\text{pat}} \to \infty} \frac{I_{\text{pat}}(t_p)}{I_d(t_p)} \approx \frac{1 - \exp(-2\beta_{\text{pat}})}{2\beta_{\text{pat}}},
\] (28)
in which the above condition that $\tau_{\text{pat}} \approx \tau_{\text{pat}}/2$ is assumed. It should be noted that the effects of homogeneous decay can be canceled by introduction of an exponential ramp, $\exp(-4\tau/T_2)$, into the pattern pulse. This would allow $\tau_{\text{pat}} \geq T_2$. However, this would require higher complexity in the input devices as well as confidence in the consistency of the homogeneous decay rate and in the shape of the decay. For now, assume that the pattern pulse is uncorrected for homogeneous decay.

**Gain Efficiency**

In an optically thick medium ($aL$ approaching or exceeding unity), the term $(aL)^2$ in approximation (26) must be replaced by the gain efficiency of the material, $\eta_{\text{ad}}$, which is a nonlinear function of $aL$. For optically thin samples ($aL \ll 1$), $\eta_{\text{ad}} = (aL)^2$. For optically thick samples, the optimal gain efficiency is a balance between the number of radiating absorbers and the absorption of the input pulses and output signal. Consider the case of a stimulated photon echo in which all three excitation pulses are assumed to undergo linear absorption and the polarization is assumed to be proportional to the cube of the excitation pulse electric fields. It has previously been shown that
\[
\eta_{\text{ad}} = [1 - \exp(-aL)]^2 \exp(-aL).
\] (29)
The assumption that all the pulses undergo linear absorption slightly underestimates the gain efficiency as one of the excitation pulses is generally a $\pi/2$ pulse that saturates the medium, and the absorber’s response to the excitation pulses is not linear but sinusoidal in nature.

**Beam-Overlap Efficiency**

From approximation (26), and Eqs. (28) and (29), an expression for the peak output intensity efficiency, $\eta_{\text{peak}}$, is obtained:
\[
\eta_{\text{peak}} \equiv \frac{I_{\text{pat}}(t_p)}{I_d(t_p)} = \frac{\eta_{\text{gates}}}{16\gamma_{\text{pat}}} \frac{\eta_{\text{decay}}}{\eta_{\text{peak}}},
\] (30)
where it is assumed that $\theta_2 = \pi/2$ and that $I_d(t_p)$ equals its maximum allowable value. For Gaussian beams, this peak efficiency is obtained only near the center of the beams. As the intensity decreases on the wings of the Gaussian, so does the efficiency. If the spatial intensity profile of the $j$th pulse can be written as
\[
I_j(r) = I_{j0}\exp\left[-(r/\sigma)^2\right],
\] (31)
where $r$ is the radial distance from the center of the Gaussian beam, then the integrated power of the $j$th pulse, $P_j$, is $\pi \sigma^2 I_{j0}$. The output intensity as a function of radial position for a coherent transient, in which the second pulse is a brief pulse and the first and third pulses are in the linear regime, is given by
\[
I_d(r) = \eta_{\text{peak}} I_{d0}(r) I_{\text{pat}}(r) \sin^2 \theta_2.
\] (32)
The total output signal power is the spatial integral of $I_d(r)$. Define the overlap efficiency $\eta_{\text{overlap}}$ as the ratio of the actual output signal power $P_d$ to the ideal output signal power, which is $\eta_{\text{peak}} P_d$ (i.e., $P_d = \eta_{\text{overlap}} \eta_{\text{peak}} P_d$). Efficiencies approaching $\eta_{\text{overlap}} = 1$ can be obtained if variable beam widths or non-diffraction-limited top-hat beams are used. However, these techniques lead to increased spatial cross talk or increased diffraction and thus lower pattern densities. In this paper, only Gaussian beams of equal width are considered. The overlap efficiency is then
\[
\eta_{\text{overlap}} = \frac{2}{\sigma^2} \int_0^\infty r^2 dr I_d(r) I_{\text{pat}}(r) \sin^2 \theta_2 \left[\frac{I_d(r)}{I_{\text{pat}}(r)}\right]^{1/2}.
\] (33)
Equation (33) can be numerically integrated. For $\theta_2 = \pi/2$, the result is $\eta_{\text{overlap}} = 0.43$. This compares with $\eta_{\text{overlap}} = 0.5$ if the sine term were ignored (set to 1). By a slight increase in the peak pulse area of pulse two, the overlap efficiency can be optimized to $\eta_{\text{overlap}} = 0.46$ when $\theta_2$ equals (0.59) $\pi$.

**E. Required Spot Size for Shot-Noise-Limited Detection**

Define the detection efficiency $\eta_{\text{det}}$ as the ratio of photons detected to the photons that are emitted in the output signal. Output coupling and scattering losses, as well as the detector’s quantum efficiency, are incorporated into the detection efficiency. The number of detected photons, $\rho$, in a correlation peak is roughly
\[
\rho = \left(\frac{\lambda}{hc\Delta v_d}\right) \eta_{\text{det}} \eta_{\text{overlap}} \eta_{\text{peak}} \pi \sigma^2 I_d,
\] (34)
assuming that the width of the correlation peak is roughly $1/\Delta v_d$. Assuming that $I_d$ is roughly equal to its maximum value given by Eq. (12), solving for $\sigma^2$ yields
\[
\sigma^2 = \frac{2 \Delta v_d \eta_{\text{det}} \eta_{\text{gates}}^2 \eta_{\text{overlap}}^2 \eta_{\text{decay}} \theta_2^4}{(\pi \eta_{\text{det}} \eta_{\text{gates}}^2 \eta_{\text{overlap}} \eta_{\text{peak}} \gamma_c \alpha^2)^4} \left(\frac{T_2 \Delta v_d}{\eta_{\text{ad}}/\sqrt{N}}\right).
\] (35)
F. Pattern Storage Density

The cross talk from neighboring spots when the third pulse is incident upon a given spatial location is now calculated for the cases in which the nearest neighbors have correlated and uncorrelated patterns and for the case in which a detection aperture is introduced.

Pattern Storage Density for Uncorrelated Nearest Neighbors

Consider two adjacent storage spots. Define an $x$-$y$ coordinate system such that two spots lie on the $x$ axis separated by $x_s$ with the third pulse centered at the origin. We ignore the saturation of the second pulse that occurs when writing the adjacent spot. This leads to an overestimate of the cross-talk signal by, at most $\theta_s^2$; the error is greatest in the region around the adjacent spot and is negligible in the region around the desired spot. The estimated peak power of the cross-talk signal from a signal adjacent spot is

$$P_c(x_s) = \int_{-\infty}^{+\infty} dx \int_{-\infty}^{+\infty} dy \eta_{\text{peak}} \theta_s^2 I_{\text{pat}}(x-x_s, y) I_s(x-x_s, y) I_s(0) / I_{\text{pat}}(0)$$

This calculation does not take into account interference between the cross-talk signal and the desired output signal. It assumes that the adjacent spots will have roughly orthogonal patterns stored in them and, therefore, the correlation peaks in the cross-talk signal and the desired output signal will not overlap temporally and coherently interfere. It is also assumed that all the patterns in the adjacent spots are roughly mutually orthogonal so that the peak level of the total cross-talk signal is just the peak level calculated above for a single adjacent spot. For Gaussian pulses, the ratio of the cross-talk signal to the desired signal is

$$\frac{P_c}{P_s} = \frac{\theta_s^2}{3\eta_{\text{overlap}}} \exp \left[ -2 \left( \frac{x_s}{\sigma} \right)^2 \right].$$

The maximum pattern density is the reciprocal of the unit cell area that depends on the two-dimensional packing geometry. Assuming hexagonal packing, the maximum pattern density for uncorrelated adjacent patterns is

$$D_{\text{max}} = \frac{4}{\sqrt{3}x_s^2} = \frac{8}{3\sqrt{3}\sigma^2 \ln \left( \frac{\theta_s^2}{3\eta_{\text{overlap}} P_c} \right)}.$$

The results are illustrated in trace (a) of Fig. 3 for $\theta_s = (0.59) \pi$ and $\eta_{\text{overlap}} = 0.46$. In the figure, the normalized pattern density is defined as $D_{\text{max}} \theta_s^2$.

Pattern Storage Density for Correlated Nearest Neighbors

If the patterns in the adjacent spots are correlated, the autocorrelation peaks temporally overlap the desired signal’s peaks and the interference between the cross-talk and desired output signals must be considered. The result depends significantly on the degree of correlation. The worst case is if all six adjacent spots are programmed identically to the central spot, but in phase or 180° out of phase with the desired signal. The magnitude of the change in the output signal power under these conditions is

$$|\Delta P_s(x_s)| = 12 \eta_{\text{peak}} \theta_s^2 \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy \frac{I_{\text{pat}}(x-x_s, y) I_s(x-x_s, y) I_s(0) I_s(0)}{I_{\text{pat}}(0) I_s(0)}$$

$$\times I_s(x, y).$$

The results are illustrated in trace (b) of Fig. 3. This condition is to be avoided as the achievable pattern densities are significantly lower in the case of correlated nearest neighbors compared with the case of uncorrelated nearest neighbors for equivalent requirements on the signal-to-noise ratio.

Pattern Storage Density with a Detection Aperture

The isolation between adjacent spots can be significantly improved by placing an aperture in the image plane of the output signal. Because the stored information is produced by the mixing of two Gaussian beams, its spatial extent is much less than that of the Gaussian third pulse, and a majority of the cross-talk signal comes from the region around the adjacent spots rather than from the region of the desired spot. The case in which the output signal is reimaged and centered on a square aperture of width $x_s$ is

$$D_{\text{max}} = \frac{4}{\sqrt{3}x_s^2} = \frac{8}{3\sqrt{3}\sigma^2 \ln \left( \frac{4\theta_s^2}{3\eta_{\text{overlap}}} P_s \right)}.$$

The results are illustrated in trace (c) of Fig. 3 for uncorrelated nearest neighbors and no detection aperture, correlated nearest neighbors and no detection aperture, and uncorrelated neighbors and a square detection aperture.
calculated numerically with Eq. (36) by reducing the limits on the integral to $\pm x_0/2$. The overestimate of the cross-talk signal and thus the underestimate of the density that is due to ignoring the saturation of pulse 2 are negligible under these conditions. The results are shown for the case of uncorrelated adjacent patterns in trace (c) of Fig. 3.

G. Local Heating

One concern with a continuous input beam is the effect of localized heating of the material. This heating could cause the temperature-dependent homogeneous line to broaden. Consider the case in which the sample is a rod of radius $r_0$ and length $L$. Assume that the rod is cooled only on its circumference and that the ends are effectively insulated. Assume that all the incident power, $P_d$, is dissipated in the form of heat uniformly along a cylinder of radius $\sigma$ and length $L$ centered in the rod. The problem is then axially symmetric, and the temperature rise at the center of the beam is easily calculated to be

$$\Delta T = \frac{[1 + 2 \ln(r_0/\sigma)]P_d}{4\pi L \kappa},$$

where $\kappa$ is the thermal conductivity of the material. In section 4, this effect is shown to be negligible.

4. Optimizing the Processor's Performance

Maximizing pattern storage density requires minimizing $\sigma^2$. Note that $\sigma^2$ is proportional to $(T_2 \Delta \nu_f)/(\tau_{pat} \Delta \nu_d)$. The limit on $\tau_{pat}$ is given in Eq. (27). The limit on $\Delta \nu_d$ is that it must be proportionally smaller than $\Delta \nu_f$ in order to minimize distortions that are due to bandwidth limitations. For a given maximum acceptable distortion level and assuming that $\tau_{pat}$ and $\Delta \nu_d$ are at their limits, the ratios $\Delta \nu_d/\Delta \nu_f = \theta_d$ and $\tau_{pat}/T_2 = e/2$ are fixed. Under these conditions, the pattern density is independent of both the inhomogeneous and homogeneous bandwidths of the medium and thus is independent of the time–bandwidth product, which is governed by the ratio of these bandwidths.

If the expression in Eq. (29) is used for $\eta_{det}$, then the minimum of $(\alpha/\eta_{det})$ is 5.32/$L$ and occurs at $\alpha = 0.523/L$. Thus $\sigma^2$ is inversely proportional to interaction length $L$ and is minimized when $L$ is at its maximum. The maximum path length through the sample is limited by diffraction. To determine the maximum acceptable path length through the sample, consider the intensity profile of a focused Gaussian beam:

$$I(x, y) = \frac{I(0, 0)}{1 + \left( \frac{x^2}{z_0^2} \right)^2 \exp \left( -\frac{r^2}{\sigma^2} \left[ 1 + \left( \frac{z}{z_0} \right)^2 \right] \right)},$$

where $z$ is the distance away from focus, and $z_0 = 2\pi \sigma^2/L$. At $\pm z_0/2$, the maximum intensity drop on axis is 20% and the maximum beam width is 12% greater than at the focus. If $L$ is larger than $z_0$, the beams will not remain in focus through the sample. This will lead to a loss in efficiency and increased cross talk with neighboring spots. Setting $L = z_0$, neglecting the slight divergence of the beam as it traverses the crystal, and solving for $\sigma^2$ yields

$$\sigma_{\text{min}}^2 = \left\{ \frac{(5.32)\left( 2^{\theta_2} \gamma_{\text{sat}} \gamma_c \right)}{\pi^2 \eta_{\text{det}} \eta_{\text{gate}} \gamma_{\text{overlap}} \gamma_{\text{decay}} \theta_y \eta_{\text{gating}}} \right\} \cdot \left( \frac{\lambda}{N} \right)^{1/2}.$$  

(43)

Solving for $\sigma_{\text{min}}^2$ fixes the value of the absorbing transition’s Einstein coefficient:

$$\gamma_{\text{sat}} = 4n\Delta \nu_f(0.523)/(N \lambda \sigma_{\text{min}}^2).$$

(44)

Finally, note that by increasing the absorption concentration $N$, the pattern density is increased. Unfortunately, increased absorber concentrations increases the rate of intercenter interactions and can lead to faster dephasing rates and spectral diffusion. These deleterious effects often occur for concentrations above $10^{18}$/cm$^3$. Equation (43) has been derived assuming that the choice of $N$ is independent of all other parameters. However, to obtain the maximum performance in practice, we must introduce the relationships between the maximum $N$ and other parameters (primarily the inhomogeneous and homogeneous bandwidths, the Einstein coefficient, and the gating efficiency) into approximation (35) before minimizing $\sigma^2$. These relationships are material dependent and have not been fully characterized. Thus Eq. (43) is used to estimate the maximum pattern density and is assumed to be valid if conservative values for the maximum concentrations are chosen.

Performance of an Optimized Processor. To estimate the performance of the continuous correlator with optimized material parameters, typical values for the material and system parameters must be assumed. Reasonable material parameters might be $\lambda = 800$ nm, $n = 1.5$, $\Delta \nu_f = 10$ GHz, and $\Delta \nu_d = 5$ kHz ($T_2 = 64$ $\mu$s). The concentration maximum is taken to be $10^{18}$ absorbers per cubic centimeter to avoid intercenter interactions. Assume that $\eta_{\text{det}} = 0.75$, $\gamma_{\text{overlap}} = 0.45$, $\theta_2 = 0.59 \pi$, $\epsilon = 0.1$, $\gamma_{\text{decay}} = 0.91$, and $\gamma_{\text{gating}} = \beta$. The value of $\epsilon$ requires that $\theta_2 = 0.75$ and $\beta_{\text{sat}} = 0.05$ and yields $\gamma_{\text{gating}} = 0.91$. Values for $\gamma_c$ and $\gamma_{\text{sat}}$ can be estimated by examining the spectrum of random binary sequences. For sequences up to 8192 in length, the ratio of the spectral power maximum to average value seldom exceeds 12. Set $\gamma_c = \gamma_{\text{sat}} = 16$ to take into account that the patterns and signals are not actually random. Take $\gamma_c$ to be the geometric mean of its extreme values of 1 and $\gamma_c$, i.e., $\gamma_c = 4$. In the shot-noise limit, the number of detected photons required for obtaining a bit-error rate of $10^{-9}$ is $\rho = 116$ if a 50% detection threshold is assumed.

The above input parameters yield the result $\sigma_{\text{min}} = 15$ $\mu$m. Assuming that roughly orthogonal (uncorre-
lated) adjacent patterns and a cross-talk signal-to-noise ratio of 10 are acceptable, the maximum pattern storage density without a detection aperture is $2.2 \times 10^{8}$ per cm$^2$. The data bandwidth is 5 GHz and the time-bandwidth product is 16 000. The required value of $\omega_{10}$ is 180 Hz, which corresponds to an oscillator strength of roughly $1.7 \times 10^{-6}$. The optimal path length through the crystal is 2.6 mm, and the optimal effective delay time of the excited state is 0.41 ms. The maximum input power is 13 $\mu$W, and the peak output signal power is 0.19 $\mu$W. (To obtain shot-noise-limited detection would require a cooled avalanche photodiode or coherent detection.) Assuming the sample is a cylinder of radius $r_0 = 5$ cm and has a low thermal conductivity of $\kappa = 0.1$ W/K cm, the temperature rise at the center of the beam is 0.7 mK and thus local heating effects are negligible.

The above is a reasonable estimate of the potential performance of the correlator with an optimized material, assuming reasonable values for the input parameters. Consider the effects a change in one or more of the input parameters has on the predicted optimized performance. The gating efficiency, $\eta_{\text{gate}}$, could possibly approach 100% or be limited to less than 1%, which would increase or decrease, respectively, the pattern density by a factor of 10. The value of $\gamma_e$ and $\gamma_{\text{pat}}$ could be made smaller with appropriate coding of the patterns and signals, yielding a slight increase in the pattern density. The performance is highly dependent on the required linearity. A factor of 2 reduction in $\epsilon$ and a factor of 2 increase in the cross-talk signal-to-noise ratio leads to a factor of 3.2 reduction in pattern density and a factor of 2 reduction in the time-bandwidth product.

As mentioned above, if the material’s homogeneous decay is well characterized, its effects can be canceled by appropriate tailoring of the pattern pulse. This may permit $\beta_{\text{pat}}$ to be increased from 0.05 to 1.0, resulting in a tenfold increase in the time–bandwidth product and a threefold increase in the pattern storage density.

The above examples assumed that the processor’s free parameters are reoptimized after each change in the input parameters. For nonoptimized parameters, approximation (35) yields the required spot size for any given configuration.

**5. Summary**

Modern communication, radar, and object recognition systems often rely on performing real-time convolutions of uninterrupted signal waveforms with multiple fixed-pattern waveforms. Electronic, acoustic, and fiber optic correlators are limited by low time-bandwidth product, low bandwidth, or programming difficulties. Coherent transient processing techniques have advantages over existing technologies because of their high bandwidth, large time-bandwidth product, and processing density. Coherent transients respond to the input pulses’ field amplitudes. This allows phase or amplitude encoding of the pattern and input signals. The continuous processing technique presented here has three distinct advantages over previously presented coherent transient processing techniques. First, the continuous data stream does not need to be broken up into overlapping segments that are shorter than the homogeneous decay time and then processed separately. Second, once the first two input pulses have programmed the solid, they need not be reentered. This greatly increases the maximum processing speed, permitting real-time processing of continuous data streams with multigigahertz bandwidths and with propagation delays only slightly greater than the duration of the programming pulses. Third, different patterns can be stored at multiple locations in the material and accessed randomly. Access rates of tens of microseconds enable fast reprogramming of the processor without reintroducing the programming pulses.

The authors gratefully acknowledge the assistance of Rudy L. Prater in verifying the derivations and editing the manuscript. This research was sponsored by the U.S. Air Force Office of Scientific Research under contract F49620-91-C-0088.

**References**