GUIDED-WAVE GaAs/AlGaAs FET OPTICAL MODULATOR BASED ON FREE-CARRIER-INDUCED BLEACHING

Indexing terms: Optoelectronics, Optical modulation, Modulators, Optical waveguides

The first optical modulators based on the free-carrier bleaching effect have been demonstrated. In these single-quantum-well FET optical modulator (FETOM) devices, the FET gate is self-aligned to a waveguide. They exhibit a 3 : 1 extinction ratio for a 10 V change in applied voltage to the gate electrode of a 750 μm FETOM waveguide.

Introduction: Quantum-well heterostructures offer unique physical mechanisms for efficient high-speed optical modulation in small structures. Recently, we proposed a new effect for absorption modulation in waveguiding single-quantum-well (SQW) heterostructures.[1,2] Free-carrier bleaching, which is based on electrically filling the bottom conduction-band states with free electrons, is capable of blocking the exciton and band-to-band absorption, depending on the degree of band-filling and electron temperature. Evidence for this effect has recently been detected at the few per cent level in a non-waveguiding configuration.[3] Three proposed optoelectronic modulating devices are (i) an FET optical modulator (FETOM), (ii) an optically readable memory element, and (iii) an optically switchable charge storage device. Here we present the first experimental results to show a useful level of optical modulation (factor of ~ 3) in a waveguiding FETOM structure designed to exploit free-carrier bleaching at room temperature.

Experiment: Fig. 1 illustrates the structure of the FETOM prepared for these experiments. Its key feature is an SQW which functions both as a modulation-doped FET channel

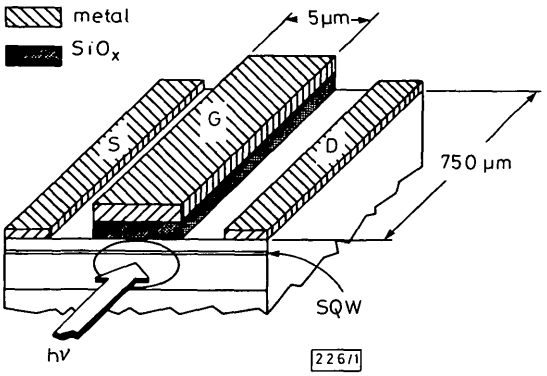


Fig. 1 *FETOM structure*

A 1000 Å SiO_x dielectric separates gate electrode from semiconductor, to promote optical waveguiding. G = gate, D = drain, S = source

and as an active absorber element embedded in an optical waveguide. The two-dimensional (2D) free-carrier density in the SQW conduction band is modulated by the gate-source voltage, just as in a conventional FET. When the channel is depleted of conduction-band carriers the SQW acts as an optical absorber for radiation tuned to the bandgap energy; filling the channel renders the SQW transparent, since absorption of a photon requires empty states in the conduction band to accept a valence-band electron. The electric fields necessary to modulate the carrier density are smaller than those causing a significant band-edge shift through the quantum-confined Stark effect. At the operating wavelength the SQW is the only optically absorbing material to be overlapped by the optical mode, and variations in its absorption result in a high extinction ratio.

In order of growth, the MBE epistructure consists of a 2 μm $Al_{0.47}Ga_{0.53}As$ (average composition) waveguide cladding layer, a 2500 Å $Al_{0.27}Ga_{0.73}As$ waveguiding layer, a 100 Å GaAs SQW, a 500 Å $Al_{0.27}Ga_{0.73}As$ waveguiding layer, and a 75 Å GaAs contact layer. Together, the two $Al_{0.27}Ga_{0.73}As$ layers support a planar single-mode waveguide containing the SQW, with the top layer containing 3.5×10^{12} cm^{-2} doping (in the 350 Å closest to the SQW) for the SQW FET channel. This doping level is chosen large enough to permit (i) low source resistance and (ii) sufficient 2D carrier concentration in the normally conducting SQW for free-carrier bleaching, while small enough so that (iii) the channel can be fully depleted by gate action, and (iv) the top $Al_{0.27}Ga_{0.73}As$ layer is depleted at equilibrium. The $Al_{0.47}Ga_{0.53}As$ cladding layer is sufficiently thick to prevent the desired optical mode from significantly coupling to substrate modes. It includes 40 Å-period superlattice buffers 500 Å-thick spaced every 0.5 μm to ensure high-quality morphology. The top GaAs layer facilitates Au–Ge/Ni alloyed ohmic contacts.

After ion-milling mesas for electrical isolation between devices, source and drain ohmic metal was deposited, patterned and alloyed to contact the SQW channel on either side of the gate. Preliminary experiments indicated a need for an optically confirming dielectric layer (refractive index $n < 3.4$) to enhance waveguiding in the vicinity of the gate. Thus, a 1000 Å layer of SiO_x and then Ti/Au were e-beam-evaporated and lifted off to form a self-aligned, 5 μm-long, metal-insulator-semiconductor (MIS) gate structure. The dielectric in conjunction with the planar waveguide creates an optically laterally confined, 5 μm-wide, strip-loaded waveguide.

The depletion-mode (i.e. normally conducting) FETOMs display good FET action (Fig. 2) with a pinch-off voltage $V_{gs} = -20$ V. The channel reaches its maximum saturation

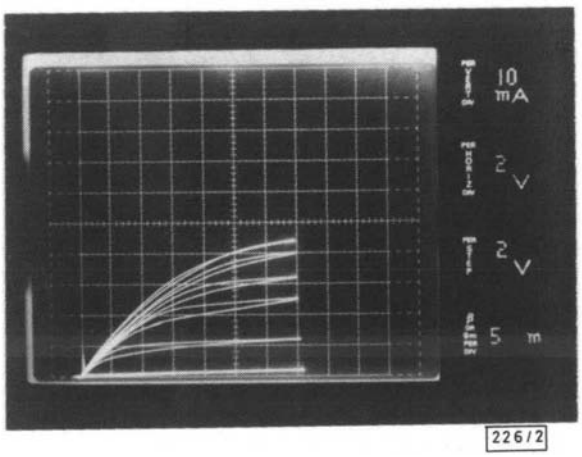


Fig. 2 *Electrical output characteristics of FETOM with 5 μm gate length (waveguide width) and 750 μm gate width (waveguide length)*

Gate biases from top to bottom trace are −10, −12, −14, −16, −18 and −20 V, respectively. Closely similar output characteristics with threshold voltage of −3.5 V were obtained in devices without the 1000 Å SiO_x layer (not shown)

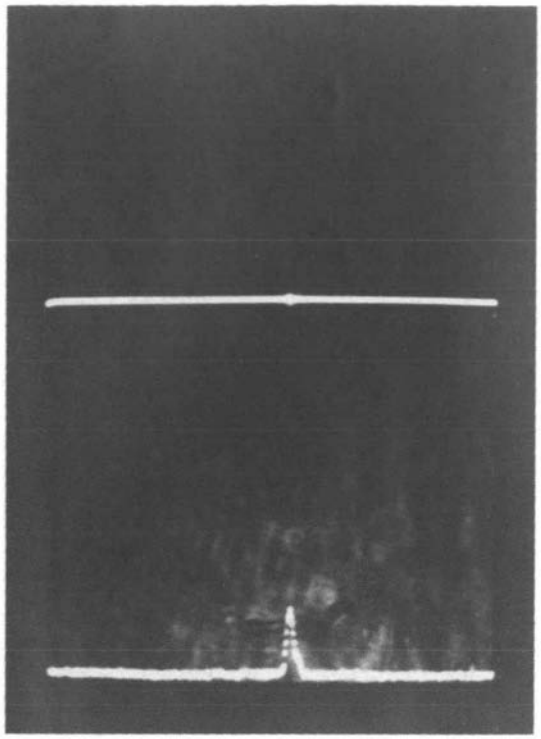
current for $V_{gs} \geq -10$ V. Owing to a 40 μm source-drain spacing, a large source resistance of $R_s = 40\,\Omega$ mm contributes to a low transconductance of 10 mS/mm, which could be improved in future designs through selective ion implantation. Similar devices fabricated on the same wafer but without the SiO_x layer pinched off at ~ -3.5 V, showing that most of the 20 V pinch-off in the present devices is dropped across the SiO_x layer and not in the semiconductor.

After cleaving, the FETOMs were placed on an electrical-optical probe stage allowing ×20 objective lens access to both cleaved surfaces and electrical needle probing to the device metallisation. Tunable radiation from a synchronously pumped dye laser (using LDS 821 dye) was focused on one facet. An optical near-field pattern from the other facet was recorded for several gate voltages V_{gs} at a wavelength of 855 nm and photographed from a vidicon display (Fig. 3). In a 750 μm-long FETOM waveguide a modulation extinction ratio of 3 : 1 was observed as V_{gs} was varied from −10 to −20 V, with the drain-source voltage maintained at 0 V to avoid any heating effects. As a function of V_{gs}, the optical modulation is directly correlated with the depletion of carriers from the SQW as revealed by source-drain current. As predicted, the transmitted intensity drops as carriers are depleted from the SQW channel with increasingly negative V_{gs}.

Optical absorption measurements as a function of wavelength in the unbiased device clearly show an absorption edge,

with the 2D exciton resonance suppressed as would be expected for estimated carrier densities of $\sim 5 \times 10^{11}\,cm^{-2}$. Calculations predict the attenuation of a 750 μm-long waveguide containing a depleted SQW to be in the order of $\sim$60 dB (using an SQW optical mode filling factor of 1·8% and absorption of $10^4\,cm^{-1}$). However, the attenuation in the present devices is less because a small fraction of the optical energy leaks around the highly absorbing gate region (in less-attenuated modes). Such leakage limits observation of a high optical extinction at the exciton peak wavelength in the direct absorption spectrum when the SQW is depleted.

To study the absorption modulation spectrum, Fig. 4 compares the absorption constant α as a function of wavelength for $V_{gs} = -20$ and $V_{gs} = 0$ V. This difference spectrum between the depleted and undepleted channel conditions shows a strong peak at the wavelength expected for the 2D exciton absorption, i.e. 855 nm (significantly shifted compared to the band-edge at 849 nm). Thus, the 2D excition absorption is the most strongly modulated spectral feature, with shorter-wavelength modulation corresponding to the band-to-band absorption quenching of a 2D Moss–Burstein effect. The decrease of absorption modulation at shorter wavelengths is more gradual than that at longer wavelengths, and is consistent with a thermal energy distribution (at $T = 300$°C, corresponding to $\Delta\lambda \sim 145$ Å) of the carrier density and thus the 2D Moss–Burstein effect.



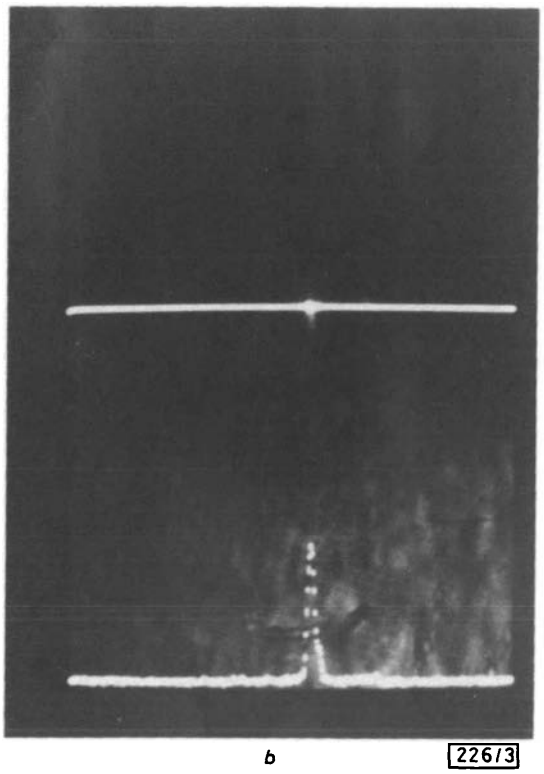


Fig. 3 *Near-field images of cleaved FETOM output face at a wavelength of 855 nm*

Bottom trace on each photograph is a plot of optical signal along horizontal line at centre

a Transmitted optical intensity for $V_{gs} = -20$ V; no carriers in quantum well (pinched-off condition). Optical absorption is high, transmitted signal is weak

b Transmitted optical intensity for $V_{gs} = -10$ V; quantum well is filled with carriers. Optical absorption is low, transmitted signal is more intense. Extinction ratio is $\sim 3:1$ after correcting for nonlinear camera response exponent γ

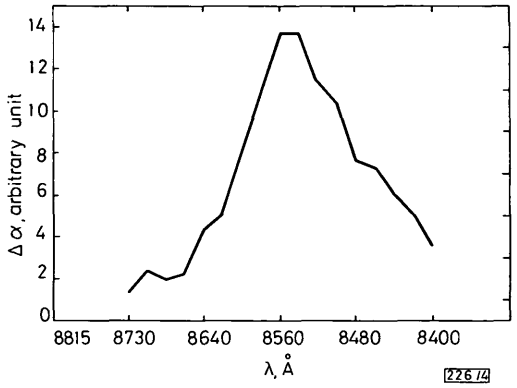


Fig. 4 *Difference in absorption constant α when SQW is filled with carriers ($V_{gs} = 0$ V) and no carriers are in SQW ($V_{gs} = -20$ V)*

Discussion: We have demonstrated experimentally the first waveguiding optical modulator based on the free-carrier bleaching effect. The correlation of transmitted optical intensity with depletion of carriers from an FET SQW channel demonstrates that the free-carrier bleaching effect is responsible for optical modulation. The difference absorption spectrum shows the suppression of an absorption peak with gate bias which cannot be explained by the quantum-confined Stark effect, since the latter would cause increased absorption for longer wavelengths at $V_{gs} = 0$ V. This would be observed as negative excursions, which are not observed in the difference spectrum of Fig. 4 at longer wavelengths.

The FETOM combines carrier transport phenomena with optical waveguiding, and is the first device to do so. Further experiments on optimised structures are indicated to study these and related effects. The use of the free-carrier density to modulate optical properties has the potential to spawn a new class of small, efficient and inherently fast optoelectronic devices.

The authors wish to thank A. C. Von Lehman and J. P. Heritage for the use of their laser.

J. H. ABELES
W. K. CHAN
A. KASTALSKY
J. P. HARBISON
L. T. FLOREZ
R. BHAT

Bell Communications Research
331 Newman Springs Road
Red Bank, NJ 07701-7020, USA

15th October 1987

References

1 KASTALSKY, A., ABELES, J. H., and LEHENY, R. F.: 'Novel optoelectronic single quantum well devices based on electron bleaching of exciton absorption', *Appl. Phys. Lett.*, 1987, **50**, pp. 708–710

2 ABELES, J. H., KASTALSKY, A., and LEHENY, R. F.: 'Novel single quantum well optoelectronic devices based on exciton bleaching', *J. Lightwave Technol.*, 1987, **LT-5**, pp. 1296–1300

3 CHEMLA, D. S., BAR-JOSEPH, I., KLINGSHIRN, C., MILLER, D. A. B., KUO, J. M., and CHANG, T. Y.: 'Optical reading of field-effect transistors by phase-space absorption quenching in a single InGaAs quantum well conducting channel', *Appl. Phys. Lett.*, 1987, **50**, pp. 585–587

REFERENCE TEMPLATE ADAPTATION IN SPEAKER-INDEPENDENT ISOLATED WORD SPEECH RECOGNITION

Indexing terms: Signal processing, Speech processing, Speech recognition

A technique which permits the adaptation of reference patterns (templates) in isolated word speech recognition systems is described. Experimental results for supervised and unsupervised adaptation with speaker-independent initial templates are presented.

Introduction: Many automatic speech recognition systems rely on a whole-word template-matching technique, using a dynamic programming algorithm[1] referred to as 'dynamic time warping' (DTW). In a system using this technique, the recognition accuracy depends critically on the relationship between the reference templates and the speech of a specific user.

There are two common approaches to deriving the reference templates for use in such a system. The speaker-dependent approach constructs reference templates by requiring the prospective user to provide at least one utterance of each word in the vocabulary, before starting to use the system. The alternative approach constructs a speaker-independent set of reference templates from utterances by a representative group of training speakers.[2] Here several templates per word are required, to allow for the variations in pronunciation among potential users.

A technique is presented here for use in template-based word recognition systems, which permits ongoing adaptation of speaker-independent reference templates during a recognition session, to take account of the information in the recognised utterances.[3–6] Adaptation can be supervised (conditional on the user's verification of each recognition) or can be unsupervised. Several different adaptation options are defined and compared here. The performance of the template adaptation technique is discussed when it is incorporated in a multiple-stage decision processor[5] for speaker-independent isolated word recognition.

System description: The structure of the recognition system has been described in detail elsewhere.[5,6] It incorporates three stages of word pattern comparison, with progressively more complex representations of the words (derived by a segmentation technique for time compression) at successive stages. This allows some templates to be eliminated with a small amount of computation: a fuller comparison is performed only for the best-matching templates. The recognition decision may be reached at any of the three stages.

When using the supervised mode for adaptation of reference templates, the system determines (by prompting the user) whether the recognition of the word is correct. If the recognition is correct, a weighted averaging procedure (with DTW alignment) is applied to adapt the correct reference template towards the recognised word. Otherwise, an incorrect reference template can be adapted away from the input word, to make recurrence of the error less likely.

In the unsupervised mode, the condition for adaptation is that the ratio of the DTW distances obtained for the best two candidate words (at the stage where the recognition decision is reached) should be greater than a specified threshold. If this condition is satisfied, the recognition is assumed to be reliable, and the weighted averaging is applied; otherwise, no adaptation takes place.

The adaptation here incorporates the 'optimisation' form of weighting:[6] the weights on the existing reference template and on the input word in the weighted averaging procedure are adjusted at successive adaptations so that the relative weight on the template increases linearly with the number of input utterances that have been used to adapt it. After any number of adaptations, each reference template is a weighted average of the initial template and the input utterances used to adapt it, in which all these inputs have equal weight.

Experimental results: Results are presented in Figs. 1 and 2 for recognition of spoken digits using speaker-independent initial templates with supervised and unsupervised adaptation, respectively. The results obtained on the same data without template adaptation are also shown, for comparison.

Speaker-independent initial templates (six templates for each digit from 'zero' to 'nine') were derived by clustering and averaging from 50 utterances of each digit, one by each of 50 training speakers. The test data consisted of three repetitions of each digit by each member of a (different) group of 49 test speakers. Because the behaviour of the system with adaptation depends on the order of the input words, the results given are averaged over four different random orderings of the 30 digits from each test speaker.

Fig. 1 shows the effects of supervised adaptation. Average recognition accuracies, with and without adaptation, are

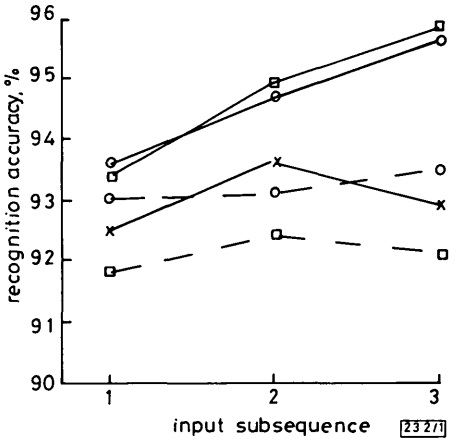


Fig. 1 *Digit recognition results with supervised adaptation*

Adaptation:
× none
□ template weight = 1·0
○ template weight = 0

Compensation:
- - - - none
——— optimal

plotted for the first, second and third 10-word subsequences of each sequence of 30 digits. Results are shown for two different weighting options. In the first weighting option (identified as 'template weight = 1·0'), when a template is adapted for the first time, the initial template and the input utterance are given equal weights. At the nth adaptation the weights on the template and on the input are adjusted to be in the ratio $n:1{\cdot}0$, so that the adapted template is the average of the initial template and all the inputs used to adapt it. In the second weighting option ('template weight = 0'), the initial template is given a weight of 0, so that the weighted averaging in the first adaptation of a template reduces to a simple replacement of the original speaker-independent template by the recognised input utterance. Here, at the nth adaptation, the weights on the template and on the input are adjusted to be in the ratio $(n-1):1{\cdot}0$. In each case, negative adaptation was applied in cases of misrecognition.

It has been shown previously[6] that recognition errors occur with template adaptation because, when some but not all of the templates have been adapted, the adapted templates correspond more closely to the speaker's voice characteristics than the unadapted ones, and so an adapted template for an incorrect candidate recognition can be closer to the input than an unadapted correct-candidate template. To compensate for this effect, an adjustment of the DTW distances has been introduced here, whereby each distance is multiplied by a quantity which increases with the number of times the template has been adapted. The results plotted with broken lines in Fig. 1